

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 8, 2004, 23:33:25 ; Search time 8113 Seconds
(without alignments)
6749.842 Million cell updates/sec

Title: US-09-998-059-1
Perfect score: 1158
Sequence: 1 cacaacatacactcaaaat.....ctctcacacacacacacacaaag 1158

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1158	100.0	1158	6	AX707049	AX707049 Sequence
2	1158	100.0	1164	6	AX707061	AX707061 Sequence
3	1158	100.0	1223	6	AX510242	AX510242 Sequence
4	1158	100.0	93695	8	ATF2009	AL021749 Arabidops
5	1158	100.0	197655	8	ATCHR169	AL161573 Arabidops
C 6	1145.4	98.9	3054	8	ATU66594	U66594 Arabidopsis
7	956.2	82.6	3100	8	ATU66916	U66916 Arabidopsis
C 8	266	23.0	329	11	BNCRU1	BNCRU1
9	198.8	17.2	4105	8	BNCRU1	BNCRU1
10	193	16.7	6049	8	AF319771	AF319771 Brassica
11	181	15.6	4901	8	RSCRUG	RSCRUG
C 12	150	13.0	150	11	BX545778	BX545778 Arabidops
C 13	119	10.3	180	11	BX296260	BX296260 Arabidops
C 14	119	10.3	205	11	BX296270	BX296270 Arabidops
C 15	81	7.0	1118	8	AY088047	AY088047 Arabidops
C 16	65.4	5.6	10326	6	AX347068	AX347068 Sequence
C 17	62	5.4	1175	8	ATU66591	U66591 Arabidopsis
C 18	59.6	5.1	110000	8	CR382131	Continuation (4 of
C 19	58	5.0	190585	2	BX572647	BX572647 Danio rer

C 20	57.4	5.0	349751	3	PFMAL4P3	AL035476 Plasmodiu
21	56.6	4.9	7236	3	AF310890	AF310890 Dictyoste
C 22	56.4	4.9	110000	2	PFMAL13_11	Continuation (12 o
C 23	56.2	4.9	6000	6	AX346035	AX346035 Sequence
24	56.2	4.9	40324	6	AX458633	AX458633 Sequence
25	55.8	4.8	5845	6	AX346564	AX346564 Sequence
26	55.8	4.8	150101	5	EX248231	EX248231 Zebrafish
27	55.4	4.8	110000	3	AC116957_1	Continuation (2 of
28	55.4	4.8	349418	3	CR382398	CR382398 Plasmodiu
29	55.2	4.8	132549	2	CR392029	CR392029 Danio rer
C 30	55	4.7	153751	3	AC116551	AC116551 Dictyoste
31	54.8	4.7	257109	3	AC116963	AC116963 Dictyoste
32	54.6	4.7	3476	6	CQ807302	CQ807302 Sequence
33	54.6	4.7	7441	6	AX348445	AX348445 Sequence
C 34	54.6	4.7	7479	6	AX339173	AX339173 Sequence
C 35	54.4	4.7	33270	3	AC116921	AC116921 Dictyoste
C 36	54.4	4.7	33203	3	CR382401	CR382401 Plasmodiu
C 37	54.2	4.7	6713	6	AX281415	AX281415 Sequence
C 38	54.2	4.7	6713	6	AX345982	AX345982 Sequence
C 39	54.2	4.7	6713	6	AX348683	AX348683 Sequence
C 40	54.2	4.7	41399	3	AC116984_5	Continuation (6 of
C 41	54	4.7	1150	8	AY093228	AY093228 Arabidops
C 42	54	4.7	126845	5	AL954683	AL954683 Zebrafish
C 43	53.8	4.6	273275	3	AE014828	AE014828 Plasmodiu
C 44	53.4	4.6	29993	6	AX825170	AX825170 Sequence
C 45	53.4	4.6	182871	3	AC117176	AC117176 Dictyoste

ALIGNMENTS

RESULT 1	AX707049	Sequence 1 from Patent WO03014347.	1158 bp	DNA	linear	PAT 04-APR-2003
AX707049	LOCUS	AX707049				
DEFINITION	AX707049	AX707049.1	GI:29563360			
ACCESSION	AX707049					
VERSION	AX707049.1					
KEYWORDS	Arabidopsis thaliana (thale cress)					
SOURCE	Arabidopsis thaliana					
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.					
REFERENCE	1	Ohlrogge,J.B., Benning,C., Gao,H., Girke,T.A. and White,J.A.				
AUTHORS	Plant seed specific promoters					
TITLE	Patent: WO 03014347-A 1 20-FEB-2003;					
JOURNAL	MICHIGAN STATE UNIVERSITY (US)					
FEATURES	Location/Qualifiers					
source	1. 1158					
	/organism="Arabidopsis thaliana"					
	/mol_type="unassigned DNA"					
	/db_xref="taxon:3702"					

ORIGIN

Query Match	100.0%;	Score 1158;	DB 6;	Length 1158;
Best Local Similarity	100.0%;	Pred. No. 4.3e-228;		
Matches 1158;	Conservative	0;	Mismatches	0;
Indels	0;	Gaps	0;	
QY	1	CACAAACATACACTCAAAATCCAGCTACATCTACTCAATTCATCTATGCACTTCATGTAA	60	
Db	1	CACAAACATACACTCAAAATCCAGCTACATCTACTCAATTCATGTAA	60	
QY	61	AACATCAAAACAGTCACAAAGTAAACAAATCAAGTCAGATTCAGCACACAAAGCCAGTAA	120	
Db	61	AACATCAAAACAGTCACAAAGTAAACAAATCAAGTCAGATTCAGCACACAAAGCCAGTAA	120	
QY	121	GATAGAAATTTAAAGAACGCTCATGCTAAGCTGCGCAAAATACCTTCCTTAATCAAAACAG	180	
Db	121	GATAGAAATTTAAAGAACGCTCATGCTAAGCTGCGCAAAATACCTTCCTTAATCAAAACAG	180	
QY	181	TAACACGAGTAAATAGCAAAATCCAGCAGAAACTCTCACCCACTCCGAAATTCAG	240	
Db	181	TAACACGAGTAAATAGCAAAATCCAGCAGAAACTCTCACCCACTCCGAAATTCAG	240	

Db 181 TAACAACGAGTAATTAGCAAAATCCGAGCAGAAAACTCTCACCACCTCCGAATTCACG 240
Qy 241 TCCTTCACTAAAATTTCCGAAGGAATCGATCAATACCAACCCCAATACACAAAATACATAA 300
Db 241 TCCTTCACTAAAATTTCCGAAGGAATCGATCAATACCAACCCCAATACACAAAATACATAA 300
Qy 301 TCAAAATGGCGAGAAATCGTACCTGGAACTTTGCTTCAAGTCGAGAGAGAGAAAGGA 360
Db 301 TCAAAATGGCGAGAAATCGTACCTGGAACTTTGCTTCAAGTCGAGAGAGAGAGAAAGGA 360
Qy 361 AGATCGTGGAGAAAGGGTTTAGGGTTTAAAGCTCAGACTTCTATTGGAGTAAATGGGACG 420
Db 361 AGATCGTGGAGAAAGGGTTTAGGGTTTAAAGCTCAGACTTCTATTGGAGTAAATGGGACG 420
Qy 421 GTGTACATATTTCCGGTTTGGAAATGAATTTGGGCTCAGCTTATGGGCTATTAGATATT 480
Db 421 GTGTACATATTTCCGGTTTGGAAATGAATTTGGGCTCAGCTTATGGGCTATTAGATATT 480
Qy 481 TGATGGGCTTCTAGTAAATACATATAAGTATTATGGGCTTAGTTTAAATAGCCCATGT 540
Db 481 TGATGGGCTTCTAGTAAATACATATAAGTATTATGGGCTTAGTTTAAATAGCCCATGT 540
Qy 541 TGGAAATATTGACACATGCTTGGCTACTAGTCTAAACATGCAACCGAACAGTGTGCG 600
Db 541 TGGAAATATTGACACATGCTTGGCTACTAGTCTAAACATGCAACCGAACAGTGTGCG 600
Qy 601 AGCAAGTCGAGCATATACAAATGGATCAACAGCGCTTAGTGTGCGCGCTCTCGCTCAT 660
Db 601 AGCAAGTCGAGCATATACAAATGGATCAACAGCGCTTAGTGTGCGCGCTCTCGCTCAT 660
Qy 661 GTGTACCTGTTTCTGCTGTTTTTTTAAATTTTCAAGTCTTTTGTGTTTATCTTCA 720
Db 661 GTGTACCTGTTTCTGCTGTTTTTTTAAATTTTCAAGTCTTTTGTGTTTATCTTCA 720
Qy 721 ATACAAATTTTGGCTGATCTTGCMAACTCTTCGATCATATCGCAATATACGTGAACA 780
Db 721 ATACAAATTTTGGCTGATCTTGCMAACTCTTCGATCATATCGCAATATACGTGAACA 780
Qy 781 CTGGTGATCTAATTTGTTGTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAG 840
Db 781 CTGGTGATCTAATTTGTTGTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAG 840
Qy 841 TGAATATATGATCATGTTTAAACATTTGATGATGATGATGATGATGATGATGATGATGAT 900
Db 841 TGAATATATGATCATGTTTAAACATTTGATGATGATGATGATGATGATGATGATGATGAT 900
Qy 901 AGTTGATGATAACGTCGAGCAAAATGAGATGATGATGATGATGATGATGATGATGATGAT 960
Db 901 AGTTGATGATAACGTCGAGCAAAATGAGATGATGATGATGATGATGATGATGATGATGAT 960
Qy 961 ACATATGCGAGAGTGAGCTACGCGCATGAAGATCAAGAGACACTTGTCTGAGCTCACAG 1020
Db 961 ACATATGCGAGAGTGAGCTACGCGCATGAAGATCAAGAGACACTTGTCTGAGCTCACAG 1020
Qy 1021 AGTGACGTGTAAGAGCTTAGACTGAAGTCCCACTGCAACCTATCTTCTGAGCTCAA 1080
Db 1021 AGTGACGTGTAAGAGCTTAGACTGAAGTCCCACTGCAACCTATCTTCTGAGCTCAA 1080
Qy 1081 ACCAGAGCTCACTTGACATATATAAACTCTCTTAAGTCCCGTCTCTTCTCATCCATCT 1140
Db 1081 ACCAGAGCTCACTTGACATATATAAACTCTCTTAAGTCCCGTCTCTTCTCATCCATCT 1140
Qy 1141 CTCACAAACAAACAAAG 1158
Db 1141 CTCACAAACAAACAAAG 1158

RESULT 2
AX707061
LOCUS AX707061 1164 bp DNA linear PAT 04-APR-2003
DEFINITION Sequence 13 from Patent WO03014347.
ACCESSION AX707061
VERSION AX707061.1 GI:29563372

KEYWORDS
SOURCE ORGANISM

Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi
1
Ohlrogge,J.B., Benning,C., Gao,H., Girke,T.A. and White,J.A.
Plant seed specific promoters
Patent: WO 03014347-A 13 20-FEB-2003;
MICHIGAN STATE UNIVERSITY (US)
FEATURES
Location/Qualifiers
1..1164
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN

Query Match 100.0%; Score 1158; DB 6; Length 1164;
Best Local Similarity 100.0%; Pred. No. 4.3e-228;
Matches 1158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CACAAACATACACTCAAAATCCAGACTCAATCTACTCAATATGCAACTTCATCATGAA 60
Db 1 CACAAACATACACTCAAAATCCAGACTCAATCTACTCAATATGCAACTTCATCATGAA 60
Qy 61 AACATCAAAACAGTCAAAATCAAAATCAAGTCAGATTTCAGACACACCAAGCCAGTAAA 120
Db 61 AACATCAAAACAGTCAAAATCAAAATCAAGTCAGATTTCAGACACACCAAGCCAGTAAA 120
Qy 121 GATAGAAAATTTAAACGAAACGCTCATGTAAAGTCGCAAAATATCTTCTATCAAAAACAG 180
Db 121 GATAGAAAATTTAAACGAAACGCTCATGTAAAGTCGCAAAATATCTTCTATCAAAAACAG 180
Qy 181 TAACAAACGAGTAAATAGCAAAATCCGAGCAGAAAATCTCTACCACTCCGAAATTCACG 240
Db 181 TAACAAACGAGTAAATAGCAAAATCCGAGCAGAAAATCTCTACCACTCCGAAATTCACG 240
Qy 241 TCCTTCACTAAAATTTTGGAAAGGAATCGATCAATACCAACCCATTCACAAATACATATA 300
Db 241 TCCTTCACTAAAATTTTGGAAAGGAATCGATCAATACCAACCCATTCACAAATACATATA 300
Qy 301 TCAAAATGGCGAGAAATCGTACCTGGAACTTTGCTTCAAGTCGAGAGAGAGAAAGGA 360
Db 301 TCAAAATGGCGAGAAATCGTACCTGGAACTTTGCTTCAAGTCGAGAGAGAGAAAGGA 360
Qy 361 AGATCGTGGAGAAAGGGTTTAGGGTTTAAAGCTCAGACTTCTATTGGAGTAAATGGGACG 420
Db 361 AGATCGTGGAGAAAGGGTTTAGGGTTTAAAGCTCAGACTTCTATTGGAGTAAATGGGACG 420
Qy 421 GTGTACATATTTCCGGTTTGGAAATGAATTTGGGCTTAGTTTAAATAGCCCATGT 480
Db 421 GTGTACATATTTCCGGTTTGGAAATGAATTTGGGCTTAGTTTAAATAGCCCATGT 480
Qy 481 TGATGGGCTTCTAGTAAATACATATAAGTATTATGGGCTTAGTTTAAATAGCCCATGT 540
Db 481 TGATGGGCTTCTAGTAAATACATATAAGTATTATGGGCTTAGTTTAAATAGCCCATGT 540
Qy 541 TGGAAATATTGACACATGCTTGGCTACTAGTCTAAACATGCAACCGAACAGTGTGCG 600
Db 541 TGGAAATATTGACACATGCTTGGCTACTAGTCTAAACATGCAACCGAACAGTGTGCG 600
Qy 601 AGCAAGTCGAGCATATACAAATGGATCAACAGCGCTTAGTGTGCGCGCTCTCGCTCAT 660
Db 601 AGCAAGTCGAGCATATACAAATGGATCAACAGCGCTTAGTGTGCGCGCTCTCGCTCAT 660
Qy 661 GTGTACCTGTTTCTGCTGTTTTTTTAAATTTTCAAGTCTTTTGTGTTTATCTTCA 720
Db 661 GTGTACCTGTTTCTGCTGTTTTTTTAAATTTTCAAGTCTTTTGTGTTTATCTTCA 720
Qy 721 ATACAAATTTTGGCTGATCTTGCMAACTCTTCGATCATATCGCAATATACGTGAACA 780
Db 721 ATACAAATTTTGGCTGATCTTGCMAACTCTTCGATCATATCGCAATATACGTGAACA 780

```
QY 781 CTGTTGATCTAATTTGTTGTTAATTTGTTAAATTTAGATTCTATTCTCCGGTTTAAAG 840
Db 781 CTGTTGATCTAATTTGTTGTTAATTTGTTAAATTTAGATTCTATTCTCCGGTTTAAAG 840
QY 841 TGAATTTATGATCATCGTTTAAACCATGTTAAAGTAAGATGATAATAAAATGATAAATTT 900
Db 841 TGAATTTATGATCATCGTTTAAACCATGTTAAAGTAAGATGATAATAAAATGATAAATTT 900
QY 901 AGTTGATGATAACCTGAAGCAAAATGAGATGAGATGATGATGATGATGATGATGATGATG 960
Db 901 AGTTGATGATAACCTGAAGCAAAATGAGATGAGATGATGATGATGATGATGATGATGATG 960
QY 961 ACATATGCGGAGAGTACGCTACGAGTGAAGATCAAGAGACATTTGCTCGAGCTCACAG 1020
Db 961 ACATATGCGGAGAGTACGCTACGAGTGAAGATCAAGAGACATTTGCTCGAGCTCACAG 1020
QY 1021 AGTACGCTGTAAGAGCTTAGACTGAAGTCCCATGCAAACTAATCTACGTGGCTCAA 1080
Db 1021 AGTACGCTGTAAGAGCTTAGACTGAAGTCCCATGCAAACTAATCTACGTGGCTCAA 1080
QY 1081 ACCACGAGCTCACCTTGACATATATAACTCTCTTAAGTCCCGTTCTCTTCATCCATCT 1140
Db 1081 ACCACGAGCTCACCTTGACATATATAACTCTCTTAAGTCCCGTTCTCTTCATCCATCT 1140
QY 1141 CTCACAAACAAACAAAAG 1158
Db 1141 CTCACAAACAAACAAAAG 1158

RESULT 3
AX510242 1223 bp DNA linear PAT 27-SEP-2002
LOCUS Sequence 4937 from Patent WO0216655.
DEFINITION AX510242
ACCESSION AX510242
VERSION AX510242.1 GI:23391479
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1
REFERENCE Harper, J.F., Krens, J., Wang, X. and Zhu, T.
AUTHORS Stress-regulated genes of plants, transgenic plants containing
TITLE same, and methods of use
JOURNAL Patent: WO 0216655-A 4937 28-FEB-2002;
The Scripps Research Institute (US) ; Syngenta Participations AG
(CH)
FEATURES
source
1. .1223
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN
Query Match 100.0%; Score 1158; DB 6; Length 1223;
Best Local Similarity 100.0%; Pred No. 4.3e-228;
Matches 1158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CACAAACATACACTCAAAATCCAGACTCACATCTACTCTCAATTATGCAACTTCATCATGAA 60
Db 63 CACAAACATACACTCAAAATCCAGACTCACATCTACTCTCAATTATGCAACTTCATCATGAA 122
QY 61 AACATCMAAAGCAGTCAAAATCAAAATCAAGTCAAGTTCAGATTCAGCACAAAGCCAGTAA 120
Db 123 AACATCMAAAGCAGTCAAAATCAAAATCAAGTCAAGTTCAGATTCAGCACAAAGCCAGTAA 182
QY 121 GATAGAAAATTTAAACGACGCTCATGCTAAAGTGGCGCAAAATCTCTTAATCAAAACAG 180
Db 183 GATAGAAAATTTAAACGACGCTCATGCTAAAGTGGCGCAAAATCTCTTAATCAAAACAG 242
QY 181 TAAACACGAGTAATTAGCAAAATCCGAGCAGAGAAAATCTCTCACCCACCTCCGAAATTCAG 240
```

RESULT 4
ATF2009
LOCUS
DEFINITION
Arabidopsis thaliana DNA chromosome 4, BAC clone F2009 (ESSA project).

```
Db 243 TAACACAGTAATTTAGCAAAATCGAGCAGAAAATCTCTACCCACCTCCGAAATTCAG 302
QY 241 TCTTCACATAAAATTTTCGAAAGGAATCGATCAATACCAACCAATTTACACAAAATACATA 300
Db 303 TCTTCACATAAAATTTTCGAAAGGAATCGATCAATACCAACCAATTTACACAAAATACATA 362
QY 301 TCAAAATGCGCAGAAATCGTACCTGGAACCTTTGCTTCAAGTCGAGAGAGAGGAAAAGGA 360
Db 363 TCAAAATGCGCAGAAATCGTACCTGGAACCTTTGCTTCAAGTCGAGAGAGAGGAAAAGGA 422
QY 361 AGATCGTGGAGAAAGGGTTTAAAGCTCAGACTTCTATTGGAGTAAATGGGACG 420
Db 423 AGATCGTGGAGAAAGGGTTTAAAGCTCAGACTTCTATTGGAGTAAATGGGACG 482
QY 421 GTGTCACATTTTCCGTTTGGAAATGAACCTTTGGGCTCAGCTTATGGCTATTAGATATT 480
Db 483 GTGTCACATTTTCCGTTTGGAAATGAACCTTTGGGCTCAGCTTATGGCTATTAGATATT 542
QY 481 TGATGGGCTTTCTAGTAAATACAAATAAAGTTATTGGGCTTAGTTTAAATAAGCCCATGT 540
Db 543 TGATGGGCTTTCTAGTAAATACAAATAAAGTTATTGGGCTTAGTTTAAATAAGCCCATGT 602
QY 541 TGGAAATATTGACACATGCTTGGCTACTAGTCTAAAACATGCAACCGAACAGTTGTCG 600
Db 603 TGGAAATATTGACACATGCTTGGCTACTAGTCTAAAACATGCAACCGAACAGTTGTCG 662
QY 601 AGACAAGTCGACATATACAAATGGAATCAACACGCTAGTGTGCGCGCTCTCGCTCAT 660
Db 663 AGACAAGTCGACATATACAAATGGAATCAACACGCTAGTGTGCGCGCTCTCGCTCAT 722
QY 661 GTGTCACTTTGTTCTCGTTTCTTTTAAATTTTCAAGTTCTTTTGTATTCTTTCA 720
Db 723 GTGTCACTTTGTTCTCGTTTCTTTTAAATTTTCAAGTTCTTTTGTATTCTTTCA 782
QY 721 ATACAAATTTTGGCTGTATCTTGCACACCTTCGATCATATCGCCAAATATAGTGAACA 780
Db 783 ATACAAATTTTGGCTGTATCTTGCACACCTTCGATCATATCGCCAAATATAGTGAACA 842
QY 781 CTGCTGATCTAAATTTGTTGTTAAATTTGTTAAATTTAGATTTCTTCTCCGGTTTAAAG 840
Db 843 CTGCTGATCTAAATTTGTTGTTAAATTTGTTAAATTTAGATTTCTTCTCCGGTTTAAAG 902
QY 841 TGAATTTATGATCATGTTTAAACCATGTTAAAGTAAGATGATAATAAAATGATAAATTT 900
Db 903 TGAATTTATGATCATGTTTAAACCATGTTAAAGTAAGATGATAATAAAATGATAAATTT 962
QY 901 AGTTGATGATACGTTGAACCAAAATGAGATGATGATGATGATGATGATGATGATGATG 960
Db 963 AGTTGATGATACGTTGAACCAAAATGAGATGATGATGATGATGATGATGATGATGATG 1022
QY 961 ACATATGCGGAGAGTACGCTACGAGTCAAGATCAAGAGACACTTGTCTCGAGCTCACAG 1020
Db 1023 ACATATGCGGAGAGTACGCTACGAGTCAAGATCAAGAGACACTTGTCTCGAGCTCACAG 1082
QY 1021 AGTACGCTGTAAGAGCTTAGACTGAAGTCCCATGCAAACTAATCTCTACGTTGCTCAA 1080
Db 1083 AGTACGCTGTAAGAGCTTAGACTGAAGTCCCATGCAAACTAATCTCTACGTTGCTCAA 1142
QY 1081 ACCACGAGCTCACCTTGACATATATAAACTCTCTTAAGTCCCGTTCTCTTCATCCATCT 1140
Db 1143 ACCACGAGCTCACCTTGACATATATAAACTCTCTTAAGTCCCGTTCTCTTCATCCATCT 1202
QY 1141 CTCACAAACAAACAAAAG 1158
Db 1203 CTCACAAACAAACAAAAG 1220
```

ATF2009 93695 bp DNA linear PLN 03-AUG-1999
Arabidopsis thaliana DNA chromosome 4, BAC clone F2009 (ESSA project).

AL021749 ACCESSION
AL021749.1 VERSION
GI:2842474 SOURCE

Arabidopsis thaliana ORGANISM
(thale cress)

Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 Bevan, M., Rose, M., Hempel, S., Entian, K.-D., Hoheisel, J.,
Mewes, H.W., Mayer, K.F.X., Lemcke, K. and Schueller, C.
Unpublished

JOURNAL
EU Arabidopsis sequencing project.
Direct submission
Submitted (30-JUL-1999) MIPS, at the Max-Planck-Institut fuer
Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:
biochem@mips.biochem.mpg.de, mayer@mips.biochem.mpg.de Project
Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge
Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,
E-mail: michael.bevan@bsrc.ac.uk

REFERENCE
Information on performance of analysis and a more detailed
annotation of this entry and other sequences of chromosomes 3, 4
and 5 can be viewed at: <http://www.mips.biochem.mpg.de/proj/thal/>.

FEATURES
source

1..93695
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/variety="Columbia"
/db_xref="taxon:3702"
/chromosome="4"
18..467
/gene="F2009.10"
complement(18..467)
/gene="F2009.10"
complement(18..467)
/gene="F2009.10"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAAL6872.1"
/db_xref="GI:2842475"
/db_xref="TREMBL:O49442"
/translation="MTSYKLFRLVIVSKRRTWACFLVVVAILLSSWSNLSINWYIG
ENHQTSGLPALVAVLLGAVFGVLMSAAALFIAPVAVIISVVVTWAFAGKSRR
VTEGRKVTEIAGVFVRVLLKEGNFVALICALLAYFFVFNYSYSSS"

exon
/gene="F2009.10"
number=1
2196..2675
/gene="F2009.20"
complement(2196..2675)
/gene="F2009.20"
complement(2196..2675)
/gene="F2009.20"
complement(2196..2675)
/codon_start=1
/product="hypothetical protein"
/protein_id="CAAL6873.2"
/db_xref="GI:5701784"
/db_xref="TREMBL:O49443"
/translation="WDKERQETWTSYKLFRVISKRTWWCLFVVVAVVAILSSRNS
LINSINWYGENHOTSSGLPALVAVLLGAVFGVLMSAAALFIAPVAVIISVVVTI
AFPVKSRKKVIEGRKVIETIAGVFVRVLLKEGNFVALICALLAYFFVFNYSYSSS"

exon
/gene="F2009.20"
number=1
4143..6188
/gene="F2009.40"
join(4143..4941,4987..5044,5096..6188)
/gene="F2009.40"
note="similarity to lectin receptor-like serine/threonine
kinase lecRK1, Arabidopsis thaliana, PIR2:S68589
Contains Protein Kinases signatures and profile: pos.
328-349 & pos. 444-456"

gene
/gene="F2009.10"
number=1
4143..4941
/gene="F2009.40"
number=1
4942..4986
/gene="F2009.40"
number=1
4987..5044
/gene="F2009.40"
number=2
5045..5095
/gene="F2009.40"
number=2
5096..6188
/gene="F2009.40"
number=3
6559..11254
/gene="F2009.30"
complement(join(6559..6669,6878..7050,7305..7428,
7537..7647,7763..7915,10629..10733,10823..10875,
11074..11254))
/gene="F2009.30"
complement(join(6559..6669,6878..7050,7305..7428,
7537..7647,7763..7915,10629..10733,10823..10875,
11074..11254))
/gene="F2009.30"
note="similarity to ENOD20 gene (small copper-binding
protein), Medicago truncatula, PATCHX:E258656"
/codon_start=1
/product="copper-binding protein-like"
/protein_id="CAAL6874.2"
/db_xref="GI:5701785"
/db_xref="GOA:O49444"
/db_xref="TREMBL:O49444"
/translation="MGLVMRFDLVLMFMVLMGLGFTISNGYKFYVGKGWVPTPSED
YSHWRNRNFQVNDLTQVPTFSLVALKTVLKELSPGPSGVPEPSDSNSRVAP
GPATPNSAGLGISTSRQLQASEEPVSLPLSGSGKEEKIKIPKRQKQAVLK
SIKQSPKKNVLAALVRGMREDALIQLQVTKRAAQVYRHHTCFVNTMLKGKVIIL
LOVIAARANAETHNEGLDRLLVIVRETTPEEAETARKLVHNFKKKSEERQLPV
HKLIETSPINRRGTKANHSSELVITILDVTCVNMEKNELDNLTNNQNIYHHNP
GP"

exon
complement(6559..6669)
/gene="F2009.30"
number=1
complement(6670..6877)
/gene="F2009.30"
number=1
complement(6878..7050)
/gene="F2009.30"
number=2
complement(7051..7304)
/gene="F2009.30"
number=2
complement(7305..7428)
/gene="F2009.30"
number=3

introns
exons
CDS

intron	complement (7429. .7536) /gene="F2009.30" /number=3	Db	64241																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																						</
--------	---	----	-------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	----

Query Match	100.0%; Score 1158; DB 8; Length 93695;
Best Local Similarity	100.0%; Pred. No. 2.7e-228;
Matches 1158; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 CACAAACATACACTCAAAATCCAGACTCACATCTACTCAATTAATGCACTTCATCATGAA
Db	64001 CACAAACATACACTCAAAATCCAGACTCACATCTACTCAATTAATGCACTTCATCATGAA
Qy	61 AACATCAAAAACAGTCAAAAGTAAACAAATCAAGTCAAGTTTCAGCACACAAAGCCAGTAAA
Db	64061 AACATCAAAAACAGTCAAAAGTAAACAAATCAAGTCAAGTTTCAGCACACAAAGCCAGTAAA
Qy	121 GATAGAAATTTAAACGAAAGCTCATCTAGCTGCGCAAAATACCTTCCCTAAATCAAAACAG
Db	64121 GATAGAAATTTAAACGAAAGCTCATCTAGCTGCGCAAAATACCTTCCCTAAATCAAAACAG
Qy	181 TAACACGAGTAAATAGCAAAATCCGAGCAGAAACTCTCACCCACCTCCGAAATTCAGG
Db	64181 TAACACGAGTAAATAGCAAAATCCGAGCAGAAACTCTCACCCACCTCCGAAATTCAGG
Qy	241 TCTTCACTAAAAATTTTCGAAAGGAATCGATCAATACCAACCCATTACACAAATACATAA

AUTHORS Sun,L. and Goodman,H.M.
 TITLE Arabidopsis genes encoding prohibitin: importance for early development
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 3054)
 AUTHORS Sun,L., Nguyen,L. and Goodman,H.M.
 TITLE Direct Submission
 JOURNAL Submitted (11-AUG-1996) Molecular Biology Department, Massachusetts General Hospital, 50 Blossom St., Boston, MA 02114, USA

FEATURES
 source 1..3054
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /strain="Columbia"
 /db_xref="taxon:3702"
 987..3054
 /gene="Atphb1"
 join(987,1048,1423,1602,1847,1963,2082,2252,2341,2553,2640,3054)
 /gene="Atphb1"
 join(1432,1602,1847,1963,2082,2252,2341,2553,2640,3054)
 /gene="Atphb1"
 /codon_start=1
 /product="prohibitin 1"
 /protein_id="AA000158.1"
 /db_xref="GI:4097694"
 /translation="MNNVKKPKIPGGGAIISLTKVGIIGGLYGLYATHSLYVVEGHR
 AINENRVLQDKVYPGTHLMPFVRYDVRPVLVSTGSDQLQWVKIGLR
 VLPRMADQLPEIYRSIGENYSEVLSIINELTKAVVAOYNASOLITQREAVSREIR
 KILTERAANFNULDVSTINLTFGKEPTAIEAKOVAQEAPKSIIVKAEQDKRS
 AVTRAQGEAKSAQLIGQAIANNQAFITLRKIEAREIAQTIANANKVYLSSDDLILN
 LQNMNLVDKNN"

ORIGIN
 Query Match 98.9%; Score 1145.4; DB 8; Length 3054;
 Best Local Similarity 99.8%; Pred. No. 1.5e-225;
 Matches 1157; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 CACAAACATACACTCAAAATCCAGACTCAGACTCAGACTCAATATGCAATTCATCATGAA 60
 Db 1369 CACAAACATACACTCAAAATCCAGACTCAGACTCAGACTCAATATGCAATTCATCATGAA 1310
 QY 61 AACATCAAAAACAGTCAAGTAAACAAATCAAGTCAGATCAGACACAAAGCCAGTAAA 120
 Db 1309 AACATCAAAAACAGTCAAGTAAACAAATCAAGTCAGATCAGACACAAAGCCAGTAAA 1250
 QY 121 GATAGAAAATTTAACGAAGCTCATGCTAAGCTGCGCAAAATCTCTCTAATCAAAAACAG 180
 Db 1249 GATAGAAAATTTAACGAAGCTCATGCTAAGCTGCGCAAAATCTCTCTAATCAAAAACAG 1190
 QY 181 TAACAACGAGTAATTTAGCAAAATCCGAGCAGAAAACTCTCACCCACCTCCGAAATTCACG 240
 Db 1189 TAACAACGAGTAATTTAGCAAAATCCGAGCAGAAAACTCTCACCCACCTCCGAAATTCACG 1130
 QY 241 TCTTCATCAAAATTTTGAAGAAATCGATCAATCAACCAACCCATTACACAAATACATAA 300
 Db 1129 TCTTCATCAAAATTTTGAAGAAATCGATCAATCAACCAACCCATTACACAAATACATAA 1070
 QY 301 TCAAAATCGCGAGAAATCGTACCTGGAACCTTTGCTTCAAGTCGCGAGAGAGAAAAGGA 360
 Db 1069 TCAAAATCGCGAGAAATCGTACCTGGAACCTTTGCTTCAAGTCGCGAGAGAGAAAAGGA 1010
 QY 361 AGATCGTGAGAAAGGGTTTGGGTTTAAAGCTTCAGACTTCATTGGAGTAATGGGAGC 420
 Db 1009 AGATCGTGAGAAAGGGTTTGGGTTTAAAGCTTCAGACTTCATTGGAGTAATGGGAGC 950
 QY 421 GTGTCAATTTCCGTTTGGAAATCACTTTGGGCTCAGCTATGGGCTATTAGATATT 480
 Db 949 GTGTCAATTTCCGTTTGGAAATCACTTTGGGCTCAGCTATGGGCTATTAGATATT 890
 QY 481 TGATGGGCTTTCTAGTAAATACAAATATAAGTTATTGGGCTTAGTTTAAATAAGCCCATGT 540

Db 889 TGATGGGCTTTCTAGTAAATACAAATATAAGTTATTGGGCTTGGTTTAAATAAGCCCATGT 830
 QY 541 TGGAAATATTTGACACATGCTTTGGCTACTAGTCTAAACATGCAACCAACAGTGTGTGCG 600
 Db 829 TGGAAATATTTGACACATGCTTTGGCTACTAGTCTAAACATGCAACCAACAGTGTGTGCG 770
 QY 601 AGACAAAGTCGACGATATACAAATGATCAACACGCTAGTGTGCGGGGTTCGGTCTCAT 660
 Db 769 AGACAAAGTCGACGATATACAAATGATCAACACGCTAGTGTGCGGGGTTCGGTCTCAT 710
 QY 661 GTGTCACTTTGTTTCTCGT-TTTTTTTTTAAATTTTCAATAGTCTTTTGTATTTCTTC 719
 Db 709 GTGTCACTTTGTTTCTCGT-TTTTTTTTTAAATTTTCAATAGTCTTTTGTATTTCTTC 650
 QY 720 AATACAAATTTTGGCTGATCTTTGCAAACTCTTCGATCATATCGCAATATACGTCGAAC 779
 Db 649 AATACAAATTTTGGCTGATCTTTGCAAACTCTTCGATCATATCGCAATATACGTCGAAC 590
 QY 780 ACTGGTGATCTAAATTTGTTGTTAAATTTGTTAAATTTAGATTTCTATCTCCGGTTTAAAA 839
 Db 589 ACTGGTGATCTAAATTTGTTGTTAAATTTGTTAAATTTAGATTTCTATCTCCGGTTTAAAA 530
 QY 840 GTGAATATATGATCATGTTTAAACATTTGTAAGTAAGATGATAATAAATGATAAATT 899
 Db 529 GTGAATATATGATCATGTTTAAACATTTGTAAGTAAGATGATAATAAATGATAAATT 470
 QY 900 TAGTTGATGATCAACGTAAGCAAAATGAGATAGATACATTTGATTTTGTTCGTATTTT 959
 Db 469 TAGTTGATGATCAACGTAAGCAAAATGAGATAGATACATTTGATTTTGTTCGTATTTT 410
 QY 960 GACATATCGCGAGAGTACGCGCATGAAGTCAAGACACTTGTCTGAGCTCACA 1019
 Db 409 GACATATCGCGAGAGTACGCGCATGAAGTCAAGACACTTGTCTGAGCTCACA 350
 QY 1020 GAGTGAGTGTAAAAGCTTAGCTGAAGTCCCATCAACCTTAATCTACGTGGTCTCA 1079
 Db 349 GAGTGAGTGTAAAAGCTTAGCTGAAGTCCCATCAACCTTAATCTACGTGGTCTCA 290
 QY 1080 AACACGAGCTCACTTCGACATATATAAACTCTCTTAAGTCCCGTTCTCTTCATCCATC 1139
 Db 289 AACACGAGCTCACTTCGACATATATAAACTCTCTTAAGTCCCGTTCTCTTCATCCATC 230
 QY 1140 TCTCACAAACAAACAAAAG 1158
 Db 229 TCTCACAAACAAACAAAAG 211

RESULT 7
 ATU66916 3100 bp DNA linear PLN 23-OCT-1996
 LOCUS Arabidopsis thaliana 12S cruciferin seed storage protein (ATCRU3)
 DEFINITION gene, complete cds.
 ACCESSION U66916
 VERSION U66916.1 GI:1628582
 KEYWORDS Arabidopsis thaliana (thale cress)
 SOURCE Arabidopsis thaliana
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 REFERENCE 1 (bases 1 to 3100)
 AUTHORS McCourt,P., Ferraioli,G. and Riggs,C.D.
 TITLE Characterization of a cruciferin-deficient mutant of Arabidopsis
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 3100)
 AUTHORS Riggs,C.D.
 TITLE Direct Submission
 JOURNAL Submitted (15-AUG-1996) Botany, University of Toronto, 1265
 Military Trail, Scarborough, Ontario M1C 1A4, Canada
 FEATURES
 source 1..3100
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"

Query Match	82.6%	Score 956.2	DB 8	Length 3100
Best Local Similarity	99.5%	Pred. No. 1.3e-186		
Matches	969	Conservative 0	Mismatches 4	Indels 1
QY	186	ACGAGTAATTTAGCAAAATCGAGCAGAGAAACTCTACCCACCTCGAAATTCACGCTCTTC		245
Db	11	ACGAGTAATTTANGTAATTCGAGCAGAGAAACTCTACCCACCTCGAAATTCACGCTCTTC		70
QY	246	ACTAAAATTTTCGAAAGGAATCGATCAATACCAACCCATTACACAAATACATATCAAA		305
Db	71	ACTAAAATTTTCGAAAGGAATCGATCAATACCAACCCATTACACAAATACATATCAAA		130
QY	306	ATGGCAGAGATCGTACTCGAAACTTTTGCTTCAAGTCGACAGAGAGAAAGGAAGATC		365
Db	131	ATGGCAGAGATCGTACTCGAAACTTTTGCTTCAAGTCGACAGAGAGAAAGGAAGATC		190
QY	366	GTGGAGAAAGGGGTTTAGGGTTTAGCTCAGACTCTATTGGAGTAAATGGGACGGTGC		425
Db	191	GTGGAGAAAGGGGTTTAGGGTTTAGCTCAGACTCTATTGGAGTAAATGGGACGGTGC		250
QY	426	ACATTTTCGGTTTGGAAATGAACTTTGGGCTCAGCTTATGGCTTATAGATATTTGATG		485
Db	251	ACATTTTCGGTTTGGAAATGAACTTTGGGCTCAGCTTATGGCTTATAGATATTTGATG		310
QY	486	GGCTTTCTAGTAATAACATTAAGTTTATGGGCTTAGTTTAAATAAGCCCATGTCGAA		545
Db	311	GGCTTTCTAGTAATAACATTAAGTTTATGGGCTTAGTTTAAATAAGCCCATGTCGAA		370
QY	546	ATATTGGACATGCTTGGCTACTAGTGTCTAAACATCAACCGAACAGTTGTGAGACA		605
Db	371	ATATTGGACATGCTTGGCTACTAGTGTCTAAACATCAACCGAACAGTTGTGAGACA		430
QY	606	AGTCGAGCATATACAAATGGATCAACACGCTAGTGTGCGCGGCTTCGCTATGTGTC		665
Db	431	AGTCGAGCATATACAAATGGATCAACACGCTAGTGTGCGCGGCTTCGCTATGTGTC		490


```

QY 960 GACATATCGGAGAGTACGCTACGGCGCATGAAGATCAAGAGACACTTCTCGAGCTCACA 1019
Db 1046 GGAGTATCGGAGATAGTCGATCGGCGCATGAACATCATGAGACACTTCTCGAGCTCACA 1105
QY 1020 GAGTACGCTGTAAGAGCTTACGATGAGTCCCAATGCAAACTTAATCTACGTGGCTCA 1079
Db 1106 GAGTGACGTGTAAGAGCATTACGACCCAGCTTCAATGCAAACTTAATCTACGTGGCTCA 1165
QY 1080 AACACAGAGCTCACTTGAATATATATAAATCTCTCTTAAG--TCCGCTTCTTCAATCCA 1137
Db 1166 AACCTTATGCTCACTCCACATATATAAATCTCTACCAAGTCTCAATGTTCTTCAATCCA 1225
QY 1138 TCTCTCAACAAACAAAA 1157
Db 1226 TCTATCAACAAACACAAA 1245

RESULT 10
AF319771
LOCUS 6049 bp DNA linear PLN 13-FEB-2001
DEFINITION Brassica napus CUC2-like protein gene, partial cds; and prohibitin
1-like protein and cruciferin subunit genes, complete cds.
ACCESSION AF319771
VERSION AF319771.1 GI:12751301
KEYWORDS Brassica napus (rape)
SOURCE Brassica napus
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 6049)
AUTHORS Scherthaler, J.P. and Robert, L.S.
TITLE Genomic clone from Brassica napus containing genes encoding a
cruciferin subunit, a prohibitin 1-like protein and the carboxy
terminus of a CUC2-like protein
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 6049)
AUTHORS Scherthaler, J.P. and Robert, L.S.
TITLE Direct Submission
JOURNAL Submitted (06-NOV-2000) Eastern Cereal and Oilseed Research Center,
Agriculture and Agri-Food Canada, 960 Carling Ave. K1N 2Z2,
Ottawa, Ontario K1A 0C6, Canada
FEATURES
source
1. .6049
/organism="Brassica napus"
/mol_type="genomic DNA"
/cultivar="Westar"
/db_xref="taxon:3708"
<54..>625
/product="CUC2-like protein"
<54..>625
/codon_start=3
/product="CUC2-like protein"
/protein_id="AAK07610.1"
/db_xref="GI:12751304"
/translation="VNDGTVRHHFHPSPKATYTGTRKPKVHTDSEVDGETRWKHTG
KTRPVLTGSRVGRYKKILVLYNYGKQKPKPTNMHGYHLGTNEEKEGELVSVKY
FYQTPROCGGSAVAATAKERPYLHGDSHLGGGRHLHLVHLNVKSGGSAVAGE
YNNIPATISNQTGIQNHVNGSQFFIP"
complement(join(1106..1300,1388..1600,1686..1856,
1998..2114,2377..>2553))
/product="prohibitin 1-like protein"
complement(join(1106..1300,1388..1600,1686..1856,
1998..2114,2377..>2553))
/codon_start=1
/product="prohibitin 1-like protein"
/protein_id="AAK07610.1"
/db_xref="GI:12751303"
/translation="WNLNKVKVPKPGGAVSALKIGIIGLGLYGATHSLYNDVG
HRAINFNLGVKQYDEGTHLMVFWFERYVYDVRAPYLVESTSGRDLQWVKIG
LRLVLRPMADQLEIYRLGENYSERVLPSIIHETLKAVVAQYNASQLITQREAVSRE
IRKILTORATNFNIALDDVSITTLTFGKEFTAIEAKQVAAQAEAKFIVEKAEQDK

```

```

RSVIRAGGEAKSAQLIQAIANNQAFITLTKIEAREIAQITIAHSANKVYLSDDL
LNLEQMLNDVNPKE"
3299..>3317
/note="G-string"
3931..3936
join(<3991..4294,4411..4727,4824..5267,5528..>5932)
/product="cruciferin subunit"
join(3991..4294,4411..4727,4824..5267,5528..5932)
/codon_start=1
/product="cruciferin subunit"
/protein_id="AAK07609.1"
/db_xref="GI:12751302"
/translation="MVKLHLVATFGALLVINGCLARQSLGVPVPPQIGNACNLNDLV
LQPTETIKSEAGRVYWDHNNPQIRCAQSVSLRIIEQGLYLPFTFSSPKISVYVQ
MGISRVPGCAETFMDSPQMOGQOGQOGQOGQOGQOGQOGQOGQOGQOGQOGQOG
EHSRVGDIATITAGSSHWIYNTGDOPLVILCLLDIANYQQLDRNPRTFRLAGNPG
GSOQOQOQNMLSGDPDQVLAQKLDVRLAQELQOQDSRGNIVKVGFPVPRP
LRQPYESEQHRHPRGPPQFODNGLEETICSMRTHENIDDPADVDYKPNLGRVTSV
SYLPIILQYIRLSATRGILQGNAMVLEPKYNNANIEILCTGQARIQVNDNGQNL
QOVKQGLVVIPOGFAYVQSHQNNFQWIFSKTNANAMVSTLAGRTSALRALPLEVLT
NAFQISLEEARIKFNTLTLLTRRRGQQLIEEIVEA"
5955..5960
polya_signal
ORIGIN
Query Match 16.7%; Score 193; DB 8; Length 6049;
Best Local Similarity 68.3%; Pred. No. 1.1e-29;
Matches 408; Conservative 0; Mismatches 160; Indels 29; Gaps 9;
QY 568 ACTAGTGCTAAACATGCAACCGCAAGTCTCGAGACAACTCGCAGATATACAAATGAT 627
Db 3399 ACTGCTCTTAAATGCAACCGCAAGTCTCGAGACAACTCGCAGATATACAAATGAT 627
QY 628 CAAACACGCTAGTGTGCGCGCTCTC-----GCTATGTCACCTTGTCTCTCTCTTT 682
Db 3458 CAAACCTCTAGTGTGCGCGCTCTCTCAAGTCTCAAGTGTCACTTTTGTCTCTTT 3517
QY 683 TTTTAAATTTTCAAGTCTCTTTGTTTATCTTCAATACAAATTTTGGCTGTATCT 742
Db 3518 CCTTTGAAT-----GGTAACTCTTCAATATATATATAGTTTGGCTGTATCT 3566
QY 743 TGCAAACTCTCGATCATATCGCA-----ATATAGTGAACACTGGTGAATCAATTTG 796
Db 3567 CCTAGACTGTTCGATCACAACCCAGATATTTAATCTGGATCTAGTGTATTAATTTG 3626
QY 797 TTGTGTTAATTTG--TAAATTTAGATTTCTATCTCGGTTTAAAGTGAATTTATCTATC 855
Db 3627 ATAGGTTAATTTGCAAAAAGTAGATTTCTCTATCCCTTTTATATATGTTTAAATGAAT 3686
QY 856 ATGGTTAAACATTTGTAAGTAAAGATGAATAATAAATGATAAATTTAGTTGATGATAACG 915
Db 3687 TAAGACGATTAAAAATGAAATTAATCA-CTAGAGAGATTCCTCTGCTGTTGATAACA 3745
QY 916 TGAAGCAAAAATGAGATAGATACATTTGATTTTGTCTGATTTTGCATATATCGGAGAT 975
Db 3746 TGAACCTAAAATGAGATACATA-TTTGAAATTTTGTCTGTTTGGAGTATCGGAGATA 3804
QY 976 GAGCTACCGCATGAAGATCAAGAGACACTTCTCGAGCTCAAGAGTCAAGCTGTAATAA 1035
Db 3805 GTGCTACGCGCATGAACATCATGAGACACTTGTCTTCAAGTCAAGAGTCAAGCTGTAATAA 3864
QY 1036 GCTTACGTAAGTCCCATGCAAACTTAATCTTACGTGGCTCAAAACCGAGCTCACTT 1095
Db 3865 CCATAGACCCACGACTTTCATGCAAACTAGTTCTTCAAGTGGCAAACTTCACTGCTCATC 3924
QY 1096 GACATATATAAATCTCTCTTAAG--TCCGCTTCTTCTCATCTCTCTCAACAA 1150
Db 3925 CAC-ATATATAAATCTCTCAAGTCTCCATGTTTCTTCTCATCTCTCTCAACAA 3980

RESULT 11
RSCRUG
LOCUS 4901 bp DNA linear PLN 06-NOV-2003
DEFINITION Raphanus sativus pgCruSes5 gene for cruciferin precursor.

```

/gene="pgCrUrS5"		15.6%;	Score 181;	DB 8;	Length 4901;
/number=3					
2015..2271					
/gene="pgCrUrS5"		63.7%;	Pred. No. 3.5e-27;		
/number=3					
2272..22676					
/gene="pgCrUrS5"					
/number=4					
2699..2704					
2829..2834					
3002..3007					
ORIGIN					
Query Match		15.6%; Score 181; DB 8; Length 4901;			
Best Local Similarity		63.7%; Pred. No. 3.5e-27;			
Matches 392; Conservative		0; Mismatches 180; Indels 43; Gaps 6;			
Qy	574	GCTAAACATGCACACCGAAGTGTGCGAGACAGTGTGCGAGACATATACAATGGAATCAACA	633		
Db	135	GCTAAACATGCACACCGAAGTGTGCGAGACATATACAATGGAATCAACA	194		
Qy	634	CGCCTAGTGTGCGCGGGTCTCGGTCTCATGTGTCACTTGTTCCTCGTGTTCCTTTTAAATTT	693		
Db	195	CTCCGATGTGCGCGGGTCTA--TCAAGTGTCACTTTTGTCTCTTTTCCTTTTGAATTT	252		
Qy	694	TTCAATAAGTTCTTTTGTTCCTTCAATACAAATTTTGGCTGTATCTTCCAAACTCTT	753		
Db	253	GGTAAC-----TTTTTTTTTGGAAACAAGTTTGGTTGTATCTCTCTAAACTGTT	302		
Qy	754	CGATCATATCGCCA-----ATATACGTGAACACTGGTGTATCTAATTTGTTGTGTTAATT	807		
Db	303	CGATTGCGAGCGCTAGATATTTTAATCTCCGACACTAGTGTGTAATTTGATTGTTAATT	362		
Qy	808	GTTAAATTTAGATTCTATTCTCGGTTTAAAAAGTGAATTAATGATCATCGTTTAAAAACA	867		
Db	363	GTCAAAAAAGTAGAATCCCATATGACCATATCTCGTTTACTCAACTACTATATGATTAA	422		
Qy	868	TTGTAAGTAAGATGATAATAA-----ATGATAAATTTAGTTGA	906		
Db	423	ATGAAATTTAGACGACTACAAATGATTACTCAATCAACACGAGTTGCTCTACTGTATAA	482		
Qy	907	TGGATAACGTGAAGCAAAAAATAGATAGATACATTTG-ATTTTGTGCTATTTTGACATA	965		
Db	483	GTGATAACATGAACCTAAAAATGATAGATACATTTGAAATTTTGTGCTGGTGTGGAGTA	542		
Qy	966	TGCGGAGGTGAGCTAGCGCATCAAGATCAAGAGACACTTGTCTCGAGCTCACAGATGA	1021		
Db	543	TGCGGAGATTGTGTCTAGCGCATCAACATTTGTGAGACACTTGTCTGAACTCACAGATGA	602		
Qy	1026	CGTGTAAAAAGCTTAGACTGAAGTCCCATGCCCATCTTAATCTCTACGTGGCTCAAAACCAC	1081		
Db	603	CGTGTAAAAACCATAGACCCACGACTTCATGCAAAACCCACTCTTACGTGGCAAAAAACCTT	662		
Qy	1086	GAGCTCACTTGCACATATATAAACTCTCTCTTAAGT---CCGTTCTCTTCATCCATCTCT	1141		
Db	663	CATGCTCACTCCACATATATAAAGCCCTACTAAGTCCACCTACTTTCTTCATCCATCTCT	722		
Qy	1143	CACAACAACAAAAA	1157		
Db	723	CACAACAACAAAAA	737		
RESULT 12					
BX545778/c					
LOCUS					
DEFINITION					
BX545778 Arabidopsis thaliana transposon insertion STS SM_3.36286, sequence tagged site.					
ACCESSION					
BX545778					
VERSION					
BX545778.1 GI:32169011					
KEYWORDS					
STS, STS, sequence tagged site.					
SOURCE					
Arabidopsis thaliana (thale cress)					
ORGANISM					
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;					

REFERENCE	AUTHORS	JOURNAL	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1	Clarke,J.H., Bowles,B., Carter,J., Hart,D., McCullagh,B., Walsh,S., Langham,S., LeGrys,C., Jones,J.D.G. and Bevan,M.	Unpublished	2	(bases 1 to 150)	Clarke,J.H.	Direct Submission	Submitted (23-JUN-2003) Clarke J.H., John Innes Centre, Colney Lane, Norwich, NR4 7UJ, UK
AT	denotes an activation tag dissociation transposon within a single line, ET an enhancer trap dissociation transposon within a trap dissociation transposon, MT a mis-expression enhancer trap dissociation transposon, SM a defective suppressor mutator transposon. 3 denotes a sequence derived from the 3' end of the transposon. 5 denotes a sequence derived from the 5' end of the transposon BSRG GARNET, ATIS project	On-line seed stock requests: http://nasc.nott.ac.uk/ NASC stock code: N122997.					
FEATURES	source						
ORIGIN							
Query Match	13.0%;	Score 150;	DB 11;	Length 150;			
Best Local Similarity	100.0%;	Pred. No. 1.2e-20;					
Matches 150;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;			
QY	895	AAATTTAGTTGATGATAACGTGAGCAAAAATGAGATAGATACATTTGCTGCT	954				
Db	150	AAATTTAGTTGATGATAACGTGAGCAAAAATGAGATAGATACATTTGCTGCT	91				
QY	955	ATTTTGACATATCGGAGAGTGAGTACGCGCATGAGATCAGAGACACATTTGCTCGAGC	1014				
Db	90	ATTTTGACATATCGGAGAGTGAGTACGCGCATGAGATCAGAGACACATTTGCTCGAGC	31				
QY	1015	TCACAGAGTGACGTGTAAAAAGCTTAGACT	1044				
Db	30	TCACAGAGTGACGTGTAAAAAGCTTAGACT	1				
RESULT 13							
LOCUS	BX296260/c						
DEFINITION	Arabidopsis thaliana transposon insertion STS SM_3.23228, sequence tagged site.						
ACCESSION	BX296260						
VERSION	BX296260.1						
KEYWORDS	STS; STS, sequence tagged site.						
SOURCE	Arabidopsis thaliana (thale cress)						
ORGANISM	Arabidopsis thaliana						
REFERENCE							
AUTHORS	Clarke,J.H., Bowles,B., Carter,J., Hart,D., McCullagh,B., Murphy,G., Latham,S., LeGrys,C., Jones,J.D.G. and Bevan,M.						
REFERENCE	Unpublished						
AUTHORS	2	(bases 1 to 180)					
TITLE	Clarke,J.H.						
JOURNAL	Direct Submission						
COMMENT	Submitted (24-MAR-2003) Clarke J.H., John Innes Centre, Colney Lane, Norwich, NR4 7UJ, UK						
AT	denotes an activation tag dissociation transposon within a single line, ET an enhancer trap dissociation transposon, GT a gene						

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 9, 2004, 07:01:07 ; Search time 816 Seconds
(without alignments)
7219.177 Million cell updates/sec

Title: US-09-998-059-1
Perfect score: 1158
Sequence: 1 cacaacatacactcaaaat.....ctctcacacacacacaaaag 1158

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3372238 seqs, 2543544081 residues

Total number of hits satisfying chosen parameters: 6744476

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1158	100.0	1158	10	US-09-998-059-1
2	1158	100.0	1164	10	US-09-998-059-13
3	1158	100.0	1223	9	US-09-938-842A-4937
4	1158	100.0	1233	11	US-09-938-842A-4937
5	65.4	5.6	10326	15	US-10-311-455-2139
6	64.2	5.5	547	17	US-10-021-323-16890
7	56.4	4.9	433	17	US-10-021-323-16890
8	56.2	4.9	6000	15	US-10-311-455-1106
9	56.2	4.9	40324	17	US-10-433-793-179
10	55.8	4.8	5945	15	US-10-311-455-1635
11	54.6	4.7	7441	16	US-10-257-186-140
12	54.6	4.7	7479	16	US-10-240-454-40
13	54.2	4.7	6713	15	US-10-311-455-1053

C 14	54.2	4.7	6713	15	US-10-240-485-79	Sequence 79, Appl
C 15	53.4	4.6	520	17	US-10-021-323-7699	Sequence 7699, Ap
C 16	53	4.6	3673778	15	US-10-312-841-1	Sequence 1, Appli
C 17	52.2	4.5	337	9	US-09-960-352-6976	Sequence 6976, Ap
C 18	51.2	4.4	529	17	US-10-021-323-10810	Sequence 10810, A
C 19	51.2	4.4	95888	11	US-09-997-722-292	Sequence 292, App
C 20	51.2	4.4	3673778	15	US-10-312-841-2	Sequence 2, Appli
C 21	51	4.4	11735	14	US-10-239-676-33	Sequence 33, Appl
C 22	51	4.4	11735	15	US-10-240-453-41	Sequence 41, Appl
C 23	51	4.4	11735	16	US-10-221-613-99	Sequence 99, Appl
C 24	51	4.4	13511	15	US-10-311-455-253	Sequence 253, App
C 25	50.6	4.4	8011	15	US-10-311-455-51	Sequence 51, Appl
C 26	50.4	4.4	419	9	US-09-960-352-11234	Sequence 11234, A
C 27	50.4	4.4	7544	14	US-10-239-676-5	Sequence 5, Appli
C 28	50.4	4.4	7544	15	US-10-240-453-13	Sequence 13, Appl
C 29	50.2	4.3	7037	15	US-10-311-455-1724	Sequence 1724, Ap
C 30	50.2	4.3	13503	15	US-10-311-455-2043	Sequence 2043, Ap
C 31	50	4.3	560	17	US-10-021-323-2253	Sequence 2253, Ap
C 32	50	4.3	5241	16	US-10-221-613-414	Sequence 414, App
C 33	49.8	4.3	6227	15	US-10-311-455-1559	Sequence 1559, Ap
C 34	49.8	4.3	6227	16	US-10-240-589C-69	Sequence 69, Appl
C 35	49.8	4.3	6476	15	US-10-311-455-514	Sequence 514, Appl
C 36	49.8	4.3	3673778	15	US-10-312-841-2	Sequence 2, Appli
C 37	49.6	4.3	499	17	US-10-021-323-10189	Sequence 10189, A
C 38	49.6	4.3	6104	15	US-10-311-455-269	Sequence 269, App
C 39	49.6	4.3	6104	16	US-10-221-714A-17	Sequence 17, Appl
C 40	49.6	4.3	83391	17	US-10-433-793-123	Sequence 123, App
C 41	49.4	4.3	7238	15	US-10-311-455-421	Sequence 421, App
C 42	49.4	4.3	17142	14	US-10-239-676-205	Sequence 205, App
C 43	49.4	4.3	17142	15	US-10-311-455-2079	Sequence 2079, Ap
C 44	49.4	4.3	17142	15	US-10-240-453-303	Sequence 303, App
C 45	49.4	4.3	17934	15	US-10-311-455-1692	Sequence 1692, Ap

ALIGNMENTS

RESULT 1

US-09-998-059-1
; Sequence 1, Application US/0998059
; Publication No. US20030005485A1
; GENERAL INFORMATION:
; APPLICANT: Ohlrogge, John B.
; APPLICANT: Benning, Christoph
; APPLICANT: Gao, Hongbo
; APPLICANT: Girke, Thomas
; APPLICANT: White, Joseph A.
; TITLE OF INVENTION: Plant Seed Specific Promoters
; FILE REFERENCE: MSU-06889
; CURRENT APPLICATION NUMBER: US/09/998,059
; CURRENT FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/250,401
; PRIOR FILING DATE: 2000-12-01
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1158
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-998-059-1

Query Match	100.0%	Score 1158;	DB 10;	Length 1158;
Best Local Similarity	100.0%;	Pred. No. 1.5e-255;		
Matches 1158;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	CACAACTACTACTCAAAATCCAGACTCACATCTACTCAATATGCACTTCATCATGAA	60	
Db	1	CACAACTACTACTCAAAATCCAGACTCACATCTACTCAATATGCACTTCATCATGAA	60	
QY	61	AACATCAAAAACAGTCAAAATCAAAATCAAGTCAGATTCAGCACAAGCCAGTAAA	120	
Db	61	AACATCAAAAACAGTCAAAATCAAAATCAAGTCAGATTCAGCACAAGCCAGTAAA	120	

Thu Oct 14 09:37:08 2004

us-09-998-059-1.rnpb

```
QY 121 GATAGAAATTTTAAAGAAAGCTCTAGCTAGCTGCGCAAAATATCTTCTTAATCAAAACAG 180
DB 121 GATAGAAATTTTAAAGAAAGCTCTAGCTAGCTGCGCAAAATATCTTCTTAATCAAAACAG 180
QY 181 TAAACACGAGTATTAAGCAAAATCCGAGCAGAAATCTCTCACCACCTCCGAAATTCACG 240
DB 181 TAAACACGAGTATTAAGCAAAATCCGAGCAGAAATCTCTCACCACCTCCGAAATTCACG 240
QY 241 TCTTCACTAAATTTTGAAGAAATCGATCAATCAACCAATCAACCAATCAACCAATCAATATA 300
DB 241 TCTTCACTAAATTTTGAAGAAATCGATCAATCAACCAATCAACCAATCAACCAATCAATATA 300
QY 301 TCAAAATGGCGAGATCGTACCTGGAAATTTGCTCAAGTCGAGAGAGAGAAAGGA 360
DB 301 TCAAAATGGCGAGATCGTACCTGGAAATTTGCTCAAGTCGAGAGAGAGAAAGGA 360
QY 361 AGATCGTGGAGAAAGGGGTTAGGTTTAAAGTCAAGCTCAGACTCTTATGGAGTAAATGGGACG 420
DB 361 AGATCGTGGAGAAAGGGGTTAGGTTTAAAGTCAAGCTCAGACTCTTATGGAGTAAATGGGACG 420
QY 421 GTGTCACATTTTCCGTTTGGAAATGAACCTTTGGGCTCAGCTTATGGGCTATTAAGATTT 480
DB 421 GTGTCACATTTTCCGTTTGGAAATGAACCTTTGGGCTCAGCTTATGGGCTATTAAGATTT 480
QY 481 TGATGGGCTTTCTAGTAAATCAATATAAGTTATTTGGGCTTATGGGCTTATTAAGATTT 540
DB 481 TGATGGGCTTTCTAGTAAATCAATATAAGTTATTTGGGCTTATGGGCTTATTAAGATTT 540
QY 541 TGGAAATATTTGACACATGCTTGGCTACTAGTCTAAACATGCAACCGACAGTTGTCG 600
DB 541 TGGAAATATTTGACACATGCTTGGCTACTAGTCTAAACATGCAACCGACAGTTGTCG 600
QY 601 AGACAAGTCGAGCATATACATAGATCAACACGCTAGTGTGCGCGCTCTCGCTCAT 660
DB 601 AGACAAGTCGAGCATATACATAGATCAACACGCTAGTGTGCGCGCTCTCGCTCAT 660
QY 661 GTGTCACCTGTTCTGCTGTTTAAATTTTCAAGTTCTTTTCAAGTTCTTTTCAAGTTCTTT 720
DB 661 GTGTCACCTGTTCTGCTGTTTAAATTTTCAAGTTCTTTTCAAGTTCTTTTCAAGTTCTTT 720
QY 721 ATACAAATTTTGGCTGATCTTGCACAACTCTTCGATCATATCGCAATATACGTGAACA 780
DB 721 ATACAAATTTTGGCTGATCTTGCACAACTCTTCGATCATATCGCAATATACGTGAACA 780
QY 781 CTGGTGATCTAATTTGTTGTTTAAATTTTCAAGTTCTTTTCAAGTTCTTTTCAAGTTCTTT 840
DB 781 CTGGTGATCTAATTTGTTGTTTAAATTTTCAAGTTCTTTTCAAGTTCTTTTCAAGTTCTTT 840
QY 841 TGAATTAATGATCATGTTTAAACATTTGTAAGTAAATGATGATGATGATGATGATGATGAT 900
DB 841 TGAATTAATGATCATGTTTAAACATTTGTAAGTAAATGATGATGATGATGATGATGATGAT 900
QY 901 AGTTGATGATTAACGTTGAAGCAAAATAGATAGATATATTTGCTGATTTTG 960
DB 901 AGTTGATGATTAACGTTGAAGCAAAATAGATAGATATATTTGCTGATTTTG 960
QY 961 ACATATGCGGAGAGTGAGCTACGCGCATCAAGATCAAGAGACACTTGTCTGAGCTCACAG 1020
DB 961 ACATATGCGGAGAGTGAGCTACGCGCATCAAGATCAAGAGACACTTGTCTGAGCTCACAG 1020
QY 1021 AGTGAGCTGTAAGAGCTTAGACTGAAGTCCCATGCAACCTTAATCTACGTGGCTCAA 1080
DB 1021 AGTGAGCTGTAAGAGCTTAGACTGAAGTCCCATGCAACCTTAATCTACGTGGCTCAA 1080
QY 1081 ACCACGAGCTCACTTGCACATATATAAATCTCTCTTAAGTCCCGTTCTCTTCATCATCT 1140
DB 1081 ACCACGAGCTCACTTGCACATATATAAATCTCTCTTAAGTCCCGTTCTCTTCATCATCT 1140
QY 1141 CTCACAAACAAACAAAAG 1158
DB 1141 CTCACAAACAAACAAAAG 1158
```

```
RESULT 2
US-09-998-059-13
; Sequence 13, Application US/09998059
; Publication No. US20030005485A1
; GENERAL INFORMATION:
; APPLICANT: Ohlrogge, John B.
; APPLICANT: Benning, Christoph
; APPLICANT: Gao, Hongbo
; APPLICANT: Girke, Thomas
; APPLICANT: White, Joseph A.
; TITLE OF INVENTION: Plant Seed Specific Promoters
; FILE REFERENCE: MSU-06689
; CURRENT APPLICATION NUMBER: US/09/998,059
; CURRENT FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/250,401
; PRIOR FILING DATE: 2000-12-01
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 1164
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; US-09-998-059-13

Query Match 100.0%; Score 1158; DB 10; Length 1164;
Best Local Similarity 100.0%; Pred. No. 1.5e-255;
Matches 1158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACAAACATACACTCAAAATCCAGACTCAGACTCACTCAATTCATGCAATTCATCATGAA 60
DB 1 CACAAACATACACTCAAAATCCAGACTCAGACTCACTCAATTCATGCAATTCATCATGAA 60
QY 61 AACATCAAAACAGTCAAAAGTAAACAAATCAAGTCAGATTCAGCACACAAAGCCAGTAAA 120
DB 61 AACATCAAAACAGTCAAAAGTAAACAAATCAAGTCAGATTCAGCACACAAAGCCAGTAAA 120
QY 121 GATAGAAATTTTAAAGAAAGCTCTAGCTAGCTGCGCAAAATATCTTCTTAATCAAAACAG 180
DB 121 GATAGAAATTTTAAAGAAAGCTCTAGCTAGCTGCGCAAAATATCTTCTTAATCAAAACAG 180
QY 181 TAAACACGAGTAAATTCGAAAGGAAATCGATCAATCAACCAACCAATTCACAAATATACATA 300
DB 181 TAAACACGAGTAAATTCGAAAGGAAATCGATCAATCAACCAACCAATTCACAAATATACATA 300
QY 241 TCTTCACTAAATTTTGAAGAAATCGATCAATCAACCAACCAATTCACAAATATACATA 300
DB 241 TCTTCACTAAATTTTGAAGAAATCGATCAATCAACCAACCAATTCACAAATATACATA 300
QY 301 TCAAAATGGCGAGATCGTACCTGGAAATCTTTGCTTCAAGTCGAGAGAGAGAAAGGA 360
DB 301 TCAAAATGGCGAGATCGTACCTGGAAATCTTTGCTTCAAGTCGAGAGAGAGAAAGGA 360
QY 361 AGATCGTGGAGAAAGGGGTTAGGTTTAAAGTCAAGCTCAGACTCTTATGGAGTAAATGGGACG 420
DB 361 AGATCGTGGAGAAAGGGGTTAGGTTTAAAGTCAAGCTCAGACTCTTATGGAGTAAATGGGACG 420
QY 421 GTGTCACATTTTCCGTTTGGAAATGAACCTTTGGGCTCAGCTTATGGGCTATTAAGATTT 480
DB 421 GTGTCACATTTTCCGTTTGGAAATGAACCTTTGGGCTCAGCTTATGGGCTATTAAGATTT 480
QY 481 TGATGGGCTTTCTAGTAAATCAATATAAGTTATTTGGGCTTATGGGCTTATTAAGATTT 540
DB 481 TGATGGGCTTTCTAGTAAATCAATATAAGTTATTTGGGCTTATGGGCTTATTAAGATTT 540
QY 541 TGGAAATATTTGACACATGCTTGGCTACTAGTCTAAACATGCAACCGACAGTTGTCG 600
DB 541 TGGAAATATTTGACACATGCTTGGCTACTAGTCTAAACATGCAACCGACAGTTGTCG 600
QY 601 AGACAAGTCGAGCATATACATAGATCAACACGCTAGTGTGCGCGCTCTCGCTCAT 660
DB 601 AGACAAGTCGAGCATATACATAGATCAACACGCTAGTGTGCGCGCTCTCGCTCAT 660
QY 661 GTGTCACCTGTTCTGCTGTTTAAATTTTCAAGTTCTTTTCAAGTTCTTTTCAAGTTCTTT 720
DB 661 GTGTCACCTGTTCTGCTGTTTAAATTTTCAAGTTCTTTTCAAGTTCTTTTCAAGTTCTTT 720
```


Db 661 ||||| 661 GTGTACCTTGTTCCTCGTCTTTTAAATTTTTCATAAGTCTTTTGTATCTTCA 720
Qy 721 ATACAAATTTTGGCTGTATCTTGCAGTCTTTCAGTATATCGCAATATACGTGAACA 780
Db 721 ATACAAATTTTGGCTGTATCTTGCAGTCTTTCAGTATATCGCAATATACGTGAACA 780
Qy 781 CTGTGATCTAAATTTTGTGTAAATTTTGTAAATTTAGATTTCTATTTCTCCGGTTTAAAG 840
Db 781 CTGTGATCTAAATTTTGTGTAAATTTTGTAAATTTAGATTTCTATTTCTCCGGTTTAAAG 840
Qy 841 TGAATTTATGATCATGTGTTAAACATTTGTAAGTAAAGATGATAATAAATGATAAATTT 900
Db 841 TGAATTTATGATCATGTGTTAAACATTTGTAAGTAAAGATGATAATAAATGATAAATTT 900
Qy 901 AGTTGATGATAACGTGAAGCAAAATGAGATAGATATATTTGATTTTGTGATTTTGG 960
Db 901 AGTTGATGATAACGTGAAGCAAAATGAGATAGATATATTTGATTTTGTGATTTTGG 960
Qy 961 ACATATCGGAGAGTACGCTACGCGCATGAAGATCAAGAGACATTTGCTCGAGCTCACAG 1020
Db 961 ACATATCGGAGAGTACGCTACGCGCATGAAGATCAAGAGACATTTGCTCGAGCTCACAG 1020
Qy 1021 AGTACGTGTTAAAGCTTAGCTGAAGTCCCATGCAAACTAATCTACGTGGCTCAA 1080
Db 1021 AGTACGTGTTAAAGCTTAGCTGAAGTCCCATGCAAACTAATCTACGTGGCTCAA 1080
Qy 1081 ACCACGAGCTCACCTTGACAATATATAACTCTCTAAGTCCCGTTCTCTTCATCCATCT 1140
Db 1081 ACCACGAGCTCACCTTGACAATATATAACTCTCTAAGTCCCGTTCTCTTCATCCATCT 1140
Qy 1141 CTCACAAACAAACAAAAG 1158
Db 1141 CTCACAAACAAACAAAAG 1158

RESULT 3

US-938-842A-4937
; Sequence 4937, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Krepes, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 4937
; LENGTH: 1223
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-4937

Query Match 100.0%; Score 1158; DB 9; Length 1223;
Best Local Similarity 100.0%; Pred. No. 1.5e-255;
Matches 1158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CACAAACATACACTCAAAATCCAGACTCACATCTACTCAATTATGCAACTTCATCATGAA 60
Db 63 CACAAACATACACTCAAAATCCAGACTCACATCTACTCAATTATGCAACTTCATCATGAA 122
Qy 61 ACATCAAAACAGTCAAAAGTAAACAAATCAAGTTCAGATTCAGCACAAAGCCAGTAA 120

Db 123 AACATCAAAACAGTCAAAAGTAAACAAATCAAGTTCAGATTCAGCACACAAAGCCAGTAA 182
Qy 121 GATAGAAATTTTAAACGAAAGCTCATGCTAAGCTGCGCAAAATACCTTCTTAATCAAAACAG 180
Db 183 GATAGAAATTTTAAACGAAAGCTCATGCTAAGCTGCGCAAAATACCTTCTTAATCAAAACAG 242
Qy 181 TAAACACAGATTAATTTAGCAAAATCCGAGCAGAAAACTCTCACCCACCTCCGAAATTCAG 240
Db 243 TAAACACAGATTAATTTAGCAAAATCCGAGCAGAAAACTCTCACCCACCTCCGAAATTCAG 302
Qy 241 TCTTCACTAAATTTTTCGAAAGGATTCGATCAATACCAACCCATTTACACAAATCAATA 300
Db 303 TCTTCACTAAATTTTTCGAAAGGATTCGATCAATACCAACCCATTTACACAAATCAATA 362
Qy 301 TCAAAATCGCAGAAATCGTACCTGAAACTTTTCTCAAGTCGACAGAGAGGAAAGGA 360
Db 363 TCAAAATCGCAGAAATCGTACCTGAAACTTTTCTCAAGTCGACAGAGAGGAAAGGA 422
Qy 361 AGATCGTGGAGAAAGGGTTTGGGTTTAAAGCTTCAGACTTCTATTGGAGTAAATGGGACG 420
Db 423 AGATCGTGGAGAAAGGGTTTGGGTTTAAAGCTTCAGACTTCTATTGGAGTAAATGGGACG 482
Qy 421 GTGTCACATTTTCCGTTTTCGAAATGAACTTTGGGCTCACGTTATGGGCTATTAGATATT 480
Db 483 GTGTCACATTTTCCGTTTTCGAAATGAACTTTGGGCTCACGTTATGGGCTATTAGATATT 542
Qy 481 TGAATGGGCTTTCTAGTAAATACAAATATAAGTTATTGGCTTAGTTTAAATGAAGCCCATGT 540
Db 543 TGAATGGGCTTTCTAGTAAATACAAATATAAGTTATTGGCTTAGTTTAAATGAAGCCCATGT 602
Qy 541 TGAATATTTTGACACATGCTTGGCTTAAAGTCTAAACATGCAACCGACAGTGTGCG 600
Db 603 TGAATATTTTGACACATGCTTGGCTTAAAGTCTAAACATGCAACCGACAGTGTGCG 662
Qy 601 AGACAAGTCGCAGCATATACAAATGGAATCAACACGCTAGTGTGCGCGCTCTCGCTCAT 660
Db 663 AGACAAGTCGCAGCATATACAAATGGAATCAACACGCTAGTGTGCGCGCTCTCGCTCAT 722
Qy 661 GTGTCACTTTGTTTCTCGTTTTCGTTTAAATTTTATAAGTCTCTTTTGTATTTTCTTCA 720
Db 723 GTGTCACTTTGTTTCTCGTTTTCGTTTAAATTTTATAAGTCTCTTTTGTATTTTCTTCA 782
Qy 721 ATACAAATTTTGGCTGTATCTTTCGAAACTCTTCGATCATATGCCAATATACGTGAACA 780
Db 783 ATACAAATTTTGGCTGTATCTTTCGAAACTCTTCGATCATATGCCAATATACGTGAACA 842
Qy 781 CTGTGATCTAAATTTGTTGTTTAAATTTGTTTAAATTTAGATTTCTATTCGCGTTTAAAG 840
Db 843 CTGTGATCTAAATTTGTTGTTTAAATTTGTTTAAATTTAGATTTCTATTCGCGTTTAAAG 902
Qy 841 TGAATTTATGATCATGTGTTAAACATTTGTAAGTAAAGTAAATTAATAAATGATAAATTT 900
Db 903 TGAATTTATGATCATGTGTTAAACATTTGTAAGTAAAGTAAATTAATAAATGATAAATTT 962
Qy 901 AGTTGATGATAACGTGAAGCAAAATGAGATAGATATATTTGATTTTGTGCTATTTTG 960
Db 963 AGTTGATGATAACGTGAAGCAAAATGAGATAGATATATTTGATTTTGTGCTATTTTG 1022
Qy 961 ACATATCGGAGAGTACGCTACGCGCATGAAGATCAAGAGACATTTGCTCGAGCTCACAG 1020
Db 1023 ACATATCGGAGAGTACGCTACGCGCATGAAGATCAAGAGACATTTGCTCGAGCTCACAG 1082
Qy 1021 AGTGACGTGTAATAAGCTTTAGACTGAAAGTCCCATGCAAACTTAATCTTACGTGGCTCAA 1080
Db 1083 AGTGACGTGTAATAAGCTTTAGACTGAAAGTCCCATGCAAACTTAATCTTACGTGGCTCAA 1142
Qy 1081 ACCACGAGCTCACTTGACAATATATAAACTCTCTAAGTCCCGTTCTCTTCATCCATCT 1140
Db 1143 ACCACGAGCTCACTTGACAATATATAAACTCTCTAAGTCCCGTTCTCTTCATCCATCT 1202
Qy 1141 CTCACAAACAAACAAAAG 1158
Db 1203 CTCACAAACAAACAAAAG 1220

Thu Oct 14 09:37:08 2004

RESULT 4
US-09-938-842A-4937
; Sequence 4937, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 4937
; LENGTH: 1223
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-4937

Query Match 100.0%; Score 1158; DB 11; Length 1223;
Best Local Similarity 100.0%; Pred. No. 1.5e-255;
Matches 1158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	CACAAACATACACTCAAAATCCAGACTCACATCTACTCAATATGCAATCTTCATCATGAA	60
Db	63	CACAAACATACACTCAAAATCCAGACTCACATCTACTCAATATGCAATCTTCATCATGAA	122
QY	61	AACATCAAAAACAGTCAAGTAACAAATCAAGTCAGATGAGTACGACACAAAGCCAGTAA	120
Db	123	AACATCAAAAACAGTCAAGTAACAAATCAAGTCAGATGAGTACGACACAAAGCCAGTAA	182
QY	121	GATAGAAATTAACGAAGCTCATGTAAGTCGGGCAAAATACCTTCTCAATCAAAACAG	180
Db	183	GATAGAAATTAACGAAGCTCATGTAAGTCGGGCAAAATACCTTCTCAATCAAAACAG	242
QY	181	TAACACAGAGTAATTAGCAAAATCCGACGAGAAATCTCACCCACCTCCGAAATTCACG	240
Db	243	TAACACAGAGTAATTAGCAAAATCCGACGAGAAATCTCACCCACCTCCGAAATTCACG	302
QY	241	TCTTCACATAAATTTTGAAGGAATCGATCAATACCAACCCATTCACCAAAATACATAA	300
Db	303	TCTTCACATAAATTTTGAAGGAATCGATCAATACCAACCCATTCACCAAAATACATAA	362
QY	301	TCBAATGCGGAGATCGTACTGGAACCTTTGCTTCAAGTCGACGAGAGAGGAAAGGA	360
Db	363	TCBAATGCGGAGATCGTACTGGAACCTTTGCTTCAAGTCGACGAGAGAGGAAAGGA	422
QY	361	AGATCGTGAGAAAGGGTTTAGGGTTTAAGTCAGACTTCTATTGGAGTAATGGGACG	420
Db	423	AGATCGTGAGAAAGGGTTTAGGGTTTAAGTCAGACTTCTATTGGAGTAATGGGACG	482
QY	421	GTGTCAATTTCCGTTTGGAAATGAACCTTTGGGCTCAGTTATGGGCTATTAGATATT	480
Db	483	GTGTCAATTTCCGTTTGGAAATGAACCTTTGGGCTCAGTTATGGGCTATTAGATATT	542
QY	481	TGATGGCTTTCTAGTAATAACATAATAGTTTATGGGCTTAGTTTAAATAAGCCCATGT	540
Db	543	TGATGGCTTTCTAGTAATAACATAATAGTTTATGGGCTTAGTTTAAATAAGCCCATGT	602
QY	541	TGGAATAATTGACATGTTTGGCTACTAGTCTTAAACATGCAACCGACAGTTGTCG	600
Db	603	TGGAATAATTGACATGTTTGGCTACTAGTCTTAAACATGCAACCGACAGTTGTCG	662

QY	601	AGACAGTCGACAGCATATACAATGGATCAAAACACGCTAGTGTGCGCGGTCTCGCTCAT	660
Db	663	AGACAGTCGACAGCATATACAATGGATCAAAACACGCTAGTGTGCGCGGTCTCGCTCAT	722
QY	661	GTGTCACTTGTCTCTCGTTTCTTTTAAATTTTTCATAAGTTCTTTTGTATCTTCA	720
Db	723	GTGTCACTTGTCTCTCGTTTCTTTTAAATTTTTCATAAGTTCTTTTGTATCTTCA	782
QY	721	ATACAAATTTTGGCTGTATCTTGCACAACTCTTCGATCATATGCCAATATACGTGAACA	780
Db	783	ATACAAATTTTGGCTGTATCTTGCACAACTCTTCGATCATATGCCAATATACGTGAACA	842
QY	781	CTGGTCACTTAATTTGTGTGTTAAATTTTAAATTTAGATTTCTTCTCGGTTTAAAG	840
Db	843	CTGGTCACTTAATTTGTGTGTTAAATTTTAAATTTAGATTTCTTCTCGGTTTAAAG	902
QY	841	TGAATTAATGATCATGTTTAAACATTTCTAGTAAGATGATAATAAAATGATAAATTT	900
Db	903	TGAATTAATGATCATGTTTAAACATTTCTAGTAAGATGATAATAAAATGATAAATTT	962
QY	901	AGTTGATGGATAACGTGAAGCAAAATGAGATAGATACATTTGTTGTTATTTTGT	960
Db	963	AGTTGATGGATAACGTGAAGCAAAATGAGATAGATACATTTGTTGTTATTTTGT	1022
QY	961	ACATATCGGAGAGTGAGCTACGCGCATGAAGATCAAGAGACACTTCTCGAGCTCACAG	1020
Db	1023	ACATATCGGAGAGTGAGCTACGCGCATGAAGATCAAGAGACACTTCTCGAGCTCACAG	1082
QY	1021	AGTGACGTGTAAGAGCTTAGACTGAAGTCCCATGCAAACTTAATCTTACGTGGCTCAA	1080
Db	1083	AGTGACGTGTAAGAGCTTAGACTGAAGTCCCATGCAAACTTAATCTTACGTGGCTCAA	1142
QY	1081	ACCAGAGCTCACCTTGACATATATAAACTCTCTAAGTCCCGTTCTTTTCATCCATCT	1140
Db	1143	ACCAGAGCTCACCTTGACATATATAAACTCTCTAAGTCCCGTTCTTTTCATCCATCT	1202
QY	1141	CTCAACAACAAACAAAAG 1158	
Db	1203	CTCAACAACAAACAAAAG 1220	

RESULT 5

US-10-311-455-2139/c
; Sequence 2139, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPNBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining the Methylation Status of Cytosine Residues in DNA
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 2139
; LENGTH: 10326
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-2139

Query Match 5.6%; Score 65.4; DB 15; Length 10326;
Best Local Similarity 48.8%; Pred. No. 0.00028;
Matches 177; Conservative 0; Mismatches 186; Indels 0; Gaps 0;

; PRIOR APPLICATION NUMBER: PCT/EP01/07470
 ; DE 10032529.7
 ; DE 10043826.1
 ; PRIOR FILING DATE: 2001-06-29
 ; 2000-06-30
 ; 2000-09-01
 ; NUMBER OF SEQ ID NOS: 178
 ; SEQ ID NO 140
 ; LENGTH: 7441
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
 US-10-257-166-140

 Query Match 4.7%; Score 54.6; DB 16; Length 7441;
 Best Local Similarity 48.8%; Pred.No. 0.071; Indels 0; Gaps 0;
 Matches 147; Conservative 0; Mismatches 154;

 QY 657 TCATGTGTACCTGTTTCCTCGTTTTTTTTTAAATTTTCATAAGTCTCTTTGTTTTATC 716
 Db 6941 TTAATTAATTTTTTGGTTTAAATTTAAATGAATTTTTTATAATTTTTTATATTGA 7000

 QY 717 TTCAATCAAAATTTTGGCTGTATCTTGCACAACTCTTCGATCATATCGCCAATATACGTG 776
 Db 7001 TTTTAAATTTTTTTGGTTTTTATTTGCGAAATTTTCGATATTATTTTTCAGGTAGTT 7060

 QY 777 AACACTGGTGATCTAAATTTGTGTGTTAAATTTGTTAAATTTAGATTTCTATCTCCGTTTA 836
 Db 7061 TATTTGGTGGCTATTAATTTTAAAGTTTTAGAGATGTTTTTATTTTTTGGATAGGTTTT 7120

 QY 837 AAAGTGAATATATGATATCATGTTAAACAACTTGAAGTAAAGTATGATATAATAATGATAA 896
 Db 7121 TGTTAGTAAATTTTGTATTTTATTTTAAATTAATTTTTTTTTTAAATATGTTAATAA 7180

 QY 897 ATTATAGTTGATGATAACGTGAACCAAAATATGATAGATACATTTGATTTTGTTCGTAT 956
 Db 7181 TTTGTTTTTGTGTAGTTTCGTTTTTAAAGAGATTTGTAATATTATATATATATATTTTTTTT 7240

 QY 957 T 957
 Db 7241 T 7241

 RESULT 12
 US-10-240-454-40
 ; Sequence 40, Application US/10240454
 ; Publication No. US20040067491A1
 ; GENERAL INFORMATION:
 ; APPLICANT: OLEK, Alexander
 ; APPLICANT: PIEPENBROCK, Christian
 ; APPLICANT: BERLIN, Kurt
 ; TITLE OF INVENTION: Diagnosis of Diseases Associated with Metabolism
 ; FILE REFERENCE: 5013.1010
 ; CURRENT APPLICATION NUMBER: US/10/240,454
 ; CURRENT FILING DATE: 2002-10-02
 ; PRIOR APPLICATION NUMBER: PCT/EP01/04016
 ; PRIOR FILING DATE: 2001-04-06
 ; PRIOR APPLICATION NUMBER: DE 10019058.8
 ; PRIOR FILING DATE: 2000-04-06
 ; PRIOR APPLICATION NUMBER: DE 10019173.8
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: DE 10032529.7
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: DE 10043826.1
 ; PRIOR FILING DATE: 2000-09-01
 ; NUMBER OF SEQ ID NOS: 68
 ; SEQ ID NO 40
 ; LENGTH: 7479
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

RESULT 14

US-10-240-485-79/c

Sequence 79, Application US/10240485

Publication No. US20030148327A1

GENERAL INFORMATION:

APPLICANT: OLEK, Alexander

APPLICANT: PIEPENBROCK, Christian

APPLICANT: BERLIN, Kurt

TITLE OF INVENTION: Diagnosis of Diseases Associated with

TITLE OF INVENTION: Metastasis

FILE REFERENCE: 5013.1007

CURRENT APPLICATION NUMBER: US/10/240,485

CURRENT FILING DATE: 2002-10-02

PRIOR APPLICATION NUMBER: PCT/EP01/03970

PRIOR FILING DATE: 2001-04-06

PRIOR APPLICATION NUMBER: DE 10019058.8

PRIOR FILING DATE: 2000-04-06

PRIOR APPLICATION NUMBER: DE 10019173.8

PRIOR FILING DATE: 2000-04-07

PRIOR APPLICATION NUMBER: DE 10032529.7

PRIOR FILING DATE: 2000-06-30

PRIOR APPLICATION NUMBER: DE 10043826.1

PRIOR FILING DATE: 2000-09-01

NUMBER OF SEQ ID NOS: 202

SEQ ID NO 79

LENGTH: 6713

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

US-10-240-485-79

Query Match 4.7%; Score 54.2; DB 15; Length 6713;

Best Local Similarity 49.1%; Pred. No. 0.084;

Matches 143; Conservative 0; Mismatches 148; Indels 0; Gaps 0;

QY 16 AAAATCCAGCTC CATCTCTCAATTATGCAACTTCATCATGAAACATCAAAAACAGT 75

DB 1045 AAAATATCCAGCTC ACTCCCAATTAATAACTTTATATAAAAACAAAACANATTACGA 986

QY 76 CAAAGTAAACAAATCAAGTCAGATTTCAGCACACAAAGCCAGTAAAGATAGAAAAATTTAAC 135

DB 985 TAACAAATATAAATCAAAAAAACCCCTCATACATACTCTAAAAATATAAATTTATTAC 926

QY 136 GAACGCTCATGCTAAGCTGCGGAAAATCTCTCTTAATCAAAACAGTAAACACGAGTAATT 195

DB 925 AACCACTATAAAAAACAAAATAAAAAATTCCTCAAAAAATAAACTAAAAACGATATAATA 866

QY 196 AGCAAAATCCGAGCAGAAAACCTCTCCACCCACCTCCGAAATTCACGCTCTTCACTAAAAATT 255

DB 865 ATCTCAACAAATCTTTAATAATATATATCTTAAAAAAAATAAATCAATACATTAATAAATA 806

QY 256 TCGAAAGGAATCGATCAATACCAACCCCAATTCACAAAAATACATAATCAAAA 306

DB 805 TTTTACTCCCATATTTTATTACACACATTTCCACATTAACATAAATAATAA 755

RESULT 15
US-10-021-323-7699/c
; Sequence 7699, Application US/10021323
; Publication NO. US20040123340A1
; GENERAL INFORMATION:

Query Match	4.6%;	Score 53.4;	DB 17;	Length 520;
Best Local Similarity	49.5%;	Pred. No. 0.039;		
Matches 138;	Conservative	0;	Mismatches 141;	Indels 0;
Gaps 0;				
662	TGTCACCTGTGTTCCCTCGATTTTTTTTTTAAATTTTTCATACAGTCTCTTTGTTTTTAFCTTCAA	721		
382	TATTAATTTTTTTTTTTTTTTTTTTTTTAAATAATATTTTTTTTTTTTTTTTTTTTTTAAAA	323		
722	TACAAATTTTGGCTGTACTCTCGAAACTCTTCGATCATATCGCCATATACGTGAACAC	781		
322	TAAATTAATTTAAAAAATTTTTTTTATTTTATTTTATCTCTTTTTTTTTTTTTTTTAAAAATTAC	263		
782	TGCTGATCTAATTTGTTGTCTTAATGTTAAAAATTTAGATTCTATTCCTCCGGTTTTAAAAAGT	841		
262	CTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAAAAATTTTACCTTTTTTTTTTTTTTTTTTAAATT	203		
842	GAATATATGTCATCGGTTTAAAAACATGTGTAAGTAAAGATGATAAAATGATAAATTTA	901		
202	TTTTTATTTTTTTTTTTTTTTTATTTAAAAATAAAAAAATAAAAAAATAATTAATAAAAAA	143		
902	GTTTGATGGAATAACGTGAAGCAAAAAAATGAGATAGATACA	940		
142	AAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA	104		

Search completed: October 9, 2004, 10:48:05
Job time : 823 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 8, 2004, 23:43:11 ; Search time 5779 Seconds
(without alignments)
7301.819 Million cell updates/sec

Title: US-09-998-059-1

Perfect score: 1158

Sequence: 1 cacaaacatacactcaaaat.....ctctcacacaacacaaaaag 1158

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hic:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gss1:*

9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	395.4	34.1	413	8 BH911706	BH911706 SALK_0716
2	333.4	28.8	400	8 BZ289690	BZ289690 SALK_0230
3	145.2	12.5	692	8 BZ084475	BZ084475 lib61d08.
4	129.4	11.2	650	8 BH555922	BH555922 BOHX116TR
C 5	83	7.2	963	9 CL515027	CL515027 SAIL_895
C 6	64	5.5	987	9 CNS014PQ	AL104456 Drosophil
7	61.6	5.3	1101	9 CNS017V2	AL108536 Drosophil
C 8	60.6	5.2	928	9 CNS00DKY	AL071865 Drosophil
9	60.4	5.2	1101	9 CNS016L1	AL06896 Drosophil
C 10	60.2	5.2	922	9 CNS00073W	AL066784 Drosophil
C 11	60	5.2	759	9 CNS06QKV	AL11257 T7 end of
C 12	59.4	5.1	324	9 CNS0243U	AL180291 Tetraodon
C 13	59.4	5.1	1187	9 CL648142	CL648142 CH213-164
C 14	59.2	5.1	1092	9 CNS020K7	AL175696 Tetraodon
C 15	58.8	5.1	1045	9 CNS03VE4	AL266197 Tetraodon
16	58.4	5.0	1001	9 CNS01400	AL103534 Drosophil
C 17	57.8	5.0	1101	9 CNS00LTF2	AL078714 Drosophil
C 18	57.6	5.0	627	1 A1525005	AL1525005 promna-4
19	57	4.9	873	9 CNS024M5	AL180950 Tetraodon
C 20	57	4.9	987	9 CNS014PQ	AL104456 Drosophil
C 21	57	4.9	1039	9 CNS006H7	AL064351 Drosophil
C 22	57	4.9	1146	9 CNS021G2	AL176843 Tetraodon
23	56.8	4.9	555	5 BQ739452	BQ739452 PfESToab4
C 24	56.8	4.9	583	9 CNS044QG	AL274417 Tetraodon

C 25	56.6	4.9	924	9 CNS07A5L	AL436159 T7 end of
C 26	56.6	4.9	982	9 CNS004WF	AL055517 Drosophil
27	56.4	4.9	764	8 CF289381	CF289381 AGENCOURT
C 28	56.4	4.9	886	6 BH177277	BH177277 008_L_22-
C 29	56.4	4.9	886	9 CNS07JUX	AL614335 T3 end of
C 30	56.4	4.9	966	9 CNS006R7	AL065822 Drosophil
C 31	56.2	4.9	892	8 AZ531076	AZ531076 ENTHX19TF
32	56.2	4.9	911	8 AZ691769	AZ691769 ENTHX19TF
C 33	56	4.8	693	9 CNS006MN	AL065759 Drosophil
C 34	56	4.8	1001	9 CNS0155H	AL105023 Drosophil
C 35	55.8	4.8	892	9 CL101621	CL101621 ISBI-38N5
36	55.8	4.8	1101	9 CNS001FB	AL060732 Drosophil
C 37	55.8	4.8	920	3 CR734152	CR734152 Tetraodon
C 38	55.6	4.8	920	9 CNS0062R	AL061710 Drosophil
39	55.4	4.8	843	9 CNS00CS1	AL059666 Drosophil
40	55.2	4.8	839	9 AG518301	AG518301 Mus muscu
41	55.2	4.8	928	9 CNS00DKY	AL071865 Drosophil
C 42	55.2	4.8	1101	9 CNS00LTF2	AL078714 Drosophil
43	55.2	4.8	1201	9 CNS0168L	AL106431 Drosophil
44	55	4.7	889	9 CNS007VQ	AL050896 Drosophil
45	54.6	4.7	1092	9 CNS020K7	AL175696 Tetraodon

ALIGNMENTS

RESULT 1
BH911706/c

LOCUS

DEFINITION

BH911706 413 bp DNA linear GSS 04-SEP-2002
SALK_071668.55.25.x Arabidopsis thaliana TDNA insertion lines
Arabidopsis thaliana genomic clone SALK_071668.55.25.x, genomic survey sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Arabidopsis thaliana (thale cress)

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE

AUTHORS

Alonso, J.M., Leisse, T.J., Barajas, P., Chen, H., Cheuk, R., Gadinab, C., Jeske, A., Karnes, M., Kim, C.J., Parker, H., Prednis, L., Shinn, P., Zimmermann, J. and Ecker, J.R.

A Sequence-Indexed Library of Insertion Mutations in the Arabidopsis Genome

Unpublished (2001)

CONTACT: Joseph R. Ecker

Salk Institute Genomic Analysis Laboratory (SIGNAL)

The Salk Institute for Biological Studies

10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

Tel: 858 453 4100 x1752

Fax: 858 558 6379

Email: ecker@salk.edu

This is single pass sequence recovered from the left border of TDNA.

Class: TDNA tagged.

Location/Qualifiers

1. 413

/organism="Arabidopsis thaliana"

/mol_type="Genomic DNA"

/ecotype="Col-0"

/db_xref="taxon:3702"

/clone="SALK_071668.55.25.x"

/note="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/tdna_protocols.html"

ORIGIN

each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/tdna_protocols.html

Query Match		28.8%;	Score 333.4;	DB 8;	Length 400;
Best Local Similarity		94.1%;	Pred. No. 5.7e-65;		
Matches 368;		Conservative 0;	Mismatches 21;	Indels 2;	Gaps 2;
QY	599	CGAGACAAGTCGCGAGCATATACAATGGATCAACACGCGCTAGTGTGCGCGCTCGCTC	658		
Db	1	CGAGACAAGTCGCGAGCATATACAATGGATCAACACGCTTATTGTGCGCGCTCGCTC	60		
QY	659	ATGTGTACCTGTGTTCTCGTGTGTTTAAATTTTTCATAAGTCTCTTTTGTATCTT	718		
Db	61	ATGTGTACCTGTGTTCTCGTGTGTTTAAATTTTTCATAAGTCTCTTTTGTATCTT	120		
QY	719	CAATACAAATTTTGGCTGTATCTTCGAACTCTTCGATCATATCGCCAATATACGTAA	778		
Db	121	CAATACAAATTTTGGCTGTATCTTCGAACTCTTCGATCATATCGCCAATATACGTAA	180		
QY	779	CACGTGTGATCTAATTTGTTGTTAATTTTAAATTTTAAATTTTAAATTTTAAAT	838		
Db	181	CACGTGTGATCTAATTTGTTGTTAATTTTAAATTTTAAATTTTAAATTTTAAAT	240		
QY	839	AGTGAATTTATGTATCATGCTTAAACATTTGTAAGTGAAGTGAATGAATGAATGAAT	898		
Db	241	AGGGAATTTATGTATCATGCTTAAACATTTGTAAGTGAAGTGAATGAATGAATGAAT	300		
QY	899	TTAGTTGATGATAACGTAAGCAAAAATTCAGATAGATACATTTGTTGTTGTTT	958		
Db	301	TTAGTTGATGATAACGTAAGCAAAAATTCAGATAGATACATTTGTTGTTGTTT	359		
QY	959	TGACATATGC-GGAGAGTGAGCTACGCGAT	988		
Db	360	TGACATATGCGGGAAGGAGCTACGCGAT	390		

RESULT 3

BZ084475

LOCUS

11b61d08.g1 B.oleracea002 Brassica oleracea genomic, genomic survey

DEFINITION

sequence.

ACCESSION

BZ084475

VERSION

BZ084475.1

GI:23714246

KEYWORDS

GSS.

SOURCE

Brassica oleracea

ORGANISM

Brassica oleracea

REFERENCE

Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R.,
Gadriab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L.,
Shinn,P., Zimmerman,J. and Ecker,J.R.
A Sequence-Indexed Library of Insertion Mutations in the
Arabidopsis Genome
Unpublished (2001)
Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGNAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: eckersalk.edu
This is single pass sequence recovered from the left border of
TDNA.
Class: TDNA tagged.

FEATURES

source

1..400

/organism="Arabidopsis thaliana"

/mol_type="genomic DNA"

/ecotype="Col-0"

/db_xref="taxon:3702"

/clone="SALK_023089.34.95.x"

/clone_lib="Arabidopsis thaliana TDNA insertion lines"

/note="PCR was performed on Arabidopsis thaliana lines"

Query Match

Best Local Similarity

Matches 402;

Conservative 0;

Mismatches 11;

Indels 0;

Gaps 0;

QY

3

CAACATACACTCAAAATCCAGACTCACATCTACTCAATTTATGCAATCTCATCATGAAA

62

Db

413

CAACATACACTCAAAATCCAGACTCACATCTACTCAATTTATGCAATCTCATCATGAAA

354

QY

63

CATCAAAAACAGTCAAAAGTAACAAAATCAAGTCAGATTCAGCACACAAAGCCAGTAAAGA

122

Db

353

CATCAAAAACAGTCAAAAGTAACAAAATCAAGTCAGATTCAGCACACAAAGCCAGTAAAGA

294

QY

123

TAGAAAATTTAAGAACGCTCATGCTAAGCTGCGGAAAATCTCTTCAATCAAAAACAGTA

182

Db

293

TAGAAAATTTAAGAACGCTCATGCTAAGCTGCGGAAAATCTCTTCAATCAAAAACAGTA

234

QY

183

ACACAGGTAATTTAGCAAAATCCGAGCAAAAATCTCTCACCCCTCCGAAAATTCAGCTC

242

Db

233

ACACAGGTAATTTAGCAAAATCCGAGCAAAAATCTCTCACCCCTCCGAAAATTCAGCTC

174

QY

243

TTCACTAAAAATTTTGGAAAGGAATCGATCAATACCAACCCATTACACAAAATACATAATC

302

Db

173

TTCACTAAAAATTTTGGAAAGGAATCGATCAATACCAACCCATTACACAAAATACATAATC

114

QY

303

AAATGCGGAGATCGTACCTGGAACCTTTGCTTCAAGTCGAGAGAGGAAAGGAAG

362

Db

113

AAATGCGGAGATCGTACCTGGAACCTTTGCTTCAAGTCGAGAGAGGAAAGGAAGTA

54

QY

363

ATCGTGAGAAAGGGTTAGGGTTTAAAGCTCAGACTTCTATTGGAGTAAATG

415

Db

53

GATTCGGAGAAAGGGTTAGGGTTTAAAGCTCAGACTTCTATTGGAGTAAATG

1

RESULT 2

BZ289690

LOCUS

SALK_023089.34.95.x Arabidopsis thaliana TDNA insertion lines

DEFINITION

Arabidopsis thaliana genomic clone SALK_023089.34.95.x, genomic survey sequence.

ACCESSION

BZ289690

VERSION

BZ289690.1

GI:24331741

KEYWORDS

GSS.

SOURCE

Arabidopsis thaliana (thale cress)

ORGANISM

Arabidopsis thaliana

REFERENCE

Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R.,
Gadriab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L.,
Shinn,P., Zimmerman,J. and Ecker,J.R.
A Sequence-Indexed Library of Insertion Mutations in the
Arabidopsis Genome
Unpublished (2001)
Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGNAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: eckersalk.edu
This is single pass sequence recovered from the left border of
TDNA.
Class: TDNA tagged.

FEATURES

source

1..400

/organism="Arabidopsis thaliana"

/mol_type="genomic DNA"

/ecotype="Col-0"

/db_xref="taxon:3702"

/clone="SALK_023089.34.95.x"

/clone_lib="Arabidopsis thaliana TDNA insertion lines"

/note="PCR was performed on Arabidopsis thaliana lines"

/notes="Vector: pOTw13; Whole genome shotgun library from flowering buds. DNA was purified from a crude nuclear prep using Brassica oleracea TO1000DH3 buds provided by Thomas Osborn at the University of Wisconsin. Genomic DNA was provided by Pablo Rabinowicz (CSHL) and the shotgun library prepared at Washington University Genome Sequencing Center."

ORIGIN

Query Match	12.5%;	Score 145.2;	DB 8;	Length 692;
Best Local Similarity	69.8%;	Pred. No. 2.8e-22;		
Matches 268;	Conservative 0;	Mismatches 108;	Indels 8;	Gaps 5;
Qy	770	ATACGTGAACACTGGTGATCTAAATTCGTGTCTTAATGTGTTAAATTTAGATTCCTATCTCT 828		
Db	195	ATAAGGGGATCTAGTGATGTAAATTTGGATACGTTAAATTTGTCAAAGAGTAGATTCCTCAT 254		
Qy	829	CCGGTTTAAAGTGAATATATATGTCATCGTTTAAACCATTTGAAGTAAAGATGATATAAA 888		
Db	255	CCCCTTTATTATATGGTTAAATGAATTTAAG--ACGATTAAAAATGAAATTAATCAACT 311		
Qy	889	AATGATATAATTTAGTTTGATGATGAACGTGAAGCAAAAATAGATGATACATTTGATTT 948		
Db	312	AGATTTCTTCTGTACTGGTTTGATACATGAACCTTAAAAATGAGAGATACAT-TTGAATTT 370		
Qy	949	TGTCGTATTTTGACATATGCGGAGGTGAGCTACGCCGATGAGATCAAGAGACACTTGC 1008		
Db	371	TGTCGTGGTTTGGAGTAGTGGGAGATAGTGTACGCCGATGACATCATGAGACACTTGC 430		
Qy	1009	TCGAGCTCAGAGTGACGTGTAAAAAGCTTTAGACTGAAGTCCCATCGAAACCTAATCC 1068		
Db	431	TTCAGCTCAGAGTGACGTGTAAAGACCATAGACCCACGACTTCATGTCAACACGAGTTC 490		
Qy	1069	TAGTGTGCTCAAAACCCAGAGCTCAGCTTGACAAATATATAAACTCCTCTAGTCCCGTTC- 1127		
Db	491	TAGTGGCAAAACCTTCATGCTCACTCCAC-ATATATAAACTCCTACCAAGTCTCCATCT 549		
Qy	1128	-TCTTCATGCATCTCTCAACAACAA 1150		
Db	550	TTCTTTCATCATCTATCACAAAA 573		

RESULT 4
 LOCUS BH655922
 DEFINITION BOHX116TR_BO_2_3_KB Brassica oleracea genomic clone BOHX116,
 650 bp DNA linear GSS 19-FEB-2002
 genomic survey sequence.
 ACCESSION BH655922
 VERSION BH655922.1 GI:18714215
 KEYWORDS GSS.
 SOURCE Brassica oleracea
 ORGANISM Brassica oleracea
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
 1 (bases 1 to 650)
 Town,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.
 Whole genome shotgun sequencing of Brassica oleracea
 Unpublished (2001)
 Contact: Chris Town
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-0208
 Email: cdtown@tigr.org
 DNA is from a doubled haploid provided by Tom Osborn.
 Seq primer: TR
 Class: sheared ends.

```

class: sheared ends.
FEATURES
  Location/Qualifiers
    1..650
      /organism="Brassica oleracea"
      /mol_type="genomic DNA"
      /strain="TO100DH3"

```

```
/db_xref="taxon:3712"  
/clone="BOHXI16"  
/clone_lib="BO_2_3_KB"  
/note="Vector: pHO51;  
genomic DNA inserted i
```

ORIGIN

Query Match	11.2%; Score 129.4; DB 8; Length 650;
Best Local Similarity	77.9%; Pred. No. 1.1e-18;
Matches 194; Conservative	0; Mismatches 51; Indels 4; Gaps 3;
QY	904 TGTGGTAACTGACGCAAAATGAGATAGATACATTTGATTTTCTCGTATTTTGACA 963
Db	9 TGGTTGATAACATGAATCTAAAAATGAGATACA-TTTGAAATTTTCTCGTGGTTGGAG 67
QY	964 TATCGCGAGAGTGAGCTACCGCATGAAGATCAAGAGACACTTGCTCGAGCTACAGAGT 1023
Db	68 TATCGCGAGATAGTGCTACCGCATGAACATCATGAGACACTTGCTTCAGCTCACAGAGT 127
QY	1024 GAGCTGTAAAAAGCTTAGACTGAAGTCCCCCATGCAAACTTAATCCTACGTGGCTCAAACC 1083
Db	128 GAGCTGTAAAGACCATAGACCCAGCACTTCATGCAAAACAGATTCCCTACGTGGCAAAACCT 187
QY	1084 ACAGGCTCACTTGACAAATATATAAACTCCCTCTTAAGTCCCGTTC--TCTTCATCCCACTCTC 1141
Db	188 TCATGCTCACTCCAC-ATATATAAATCTCTACCAAGTCTCCACTTTCTTCATCCCATCTA 246
QY	1142 TCACAACAA 1150
Db	247 TCACAAAA 255

RESULT 5
CL515027/c

LOCUS
CL515027 963 bp DNA linear GSS 01-APR-2004

DEFINITION
SAIL 895_F04.v1 SAIL Collection Arabidopsis thaliana genomic clone
SAIL 895_F04.v1, genomic survey sequence.

RESULT 5
CL515027/c

LOCUS	CL515027	963 bp	DNA	linear	GSS 01-APR-2004
DEFINITION	SAIL-895_F04.v1	SAIL Collection	Arabidopsis	thaliana	genomic clone
	SAIL-895_F04.v1	genomic survey	sequence.		

ACCESSION	CL515027
VERSION	CL515027.1
KEYWORDS	GI:46012347
SOURCE	GSS
ORGANISM	Arabidopsis thaliana (thale cress)
	Arabidopsis thaliana
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
	rosids; eurosids II; Brassicaceae; Arabidopsids.
REFERENCE	1 (bases 1 to 963)
AUTHORS	Sessions,A., Burke,E., Presting,G., Aux,G., McEliver,J., Patton,D., Dietrich,B., Ho,P., Bacwaden,J., Ko,C., Clarke,J.D., Cotton,D., Bullis,D., Snell,J., Miguel,T., Hutchison,D., Kimmerly,B., Mizel,T., Katagiri,F., Glazebrook,J., Law,M. and Goff,S.A. A high-throughput Arabidopsis reverse genetics system Plant Cell 14 (12), 2985-2994 (2002)
TITLE	
JOURNAL	

COMMENT

```

FEATURES             source
    Class: TDNA tagged.
        Location/Qualifiers
            1..963
                /organism="Arabidopsis thaliana"
                /mol_type="genomic DNA"
                /ecotype="Columbia"
                /db_xref="taxon:3702"
                /clone_lib="SAIL_F04.v1"
                /Clone_lib="SAIL_Collection"

COMMENT
    Sessions A
    Applied Trait Genetics
    Syngenta Biotechnology Inc.
    3054 Cornwallis Rd., Research Triangle Park, NC 27709, USA
    Email: allen.sessions@syngenta.com
    ABRC Stock Number CS840399; T-DNA left border flanking sequences of
    Syngenta Arabidopsis Insertion Library (SAIL) lines are available
    through the Arabidopsis Biological Resource Center (ABRC).
    Sequences represent a pool of amplified genomic regions and not
    single contiguous sequences.

```

/note="T-DNA left border sequences were isolated using a modified TAIL-PCR strategy"

```
ORIGIN
Query Match          7.2%; Score 83; DB 9; Length 963;
Best Local Similarity 57.2%; Pred. No. 3.6e-08;
Matches 191; Conservative 0; Mismatches 137; Indels 6; Gaps 3;

QY 97 GATTTCAGCACAAAGCAGTAAGATAGAGAAATTTAAACG-AACGCTCATGCTTAAGCTGC 155
Db 416 GTTTCAGCCTCATCGCCTGTCTAGATGCAACATGTGGCGCGCTCTCTGGTGAAGTGC 357

QY 156 G-CAAAATACCTCTAATCAAAACAGTAACAGAGTAATTAGCAAAATCCGAGCAGAA 214
Db 356 GCGGGAATGTTTCTTCTAAGCAGTACCATCTGGGTATGGACTATATCTAGCAATAA 297

QY 215 ACTCTCACCACCTCGAATTCACGCTTCTCACTAAAATTTTCGAAAGGAATCGATCAAT 274
Db 296 ACCTGTCGCCGTCGCAATTTCTCTTTTCAATTAATTTGGAAGTAATTTGTTCTT 237

QY 275 ACCAACCCATTACACAAAATACATAATCAAAATGGCGAATCG----TACCTGGAAC 330
Db 236 ACTNTTNCATNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 177

QY 331 TTGCTTCAAGTCGAGAGAGAGGAAAGAGATCGTGGAGAGGGGTTTAGGGTTAA 390
Db 176 GCTTGTGCGGTTCCTCCCAATGCAAAAGTAAATTTGTCAGCAACGGGTGTAGGGTTAA 117

QY 391 GCTCAGACTTCTATTGGAGTAAATGGGACGGTGT 424
Db 116 GCTCAGGCTTCCCTTGGAGTAAATGGAGACGGT 83

RESULT 6
CNS014PQ/c          987 bp DNA linear GSS 26-JUL-1999
LOCUS              Drosophila melanogaster genome survey sequence SP6 end of BAC
DEFINITION         BACN12P22 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION          AL104456
VERSION            AL104456.1 GI:5616067
KEYWORDS           Drosophila melanogaster (fruit fly)
SOURCE             Drosophila melanogaster
ORGANISM           Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE          1 (bases 1 to 987)
AUTHORS            Genoscope.
TITLE              Direct Submission
JOURNAL            Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr)
COMMENT            - Web : www.genoscope.cns.fr
Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk - This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelobAC11.

FEATURES
source            1..987
Location/Qualifiers
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACN12P22"
/clone_lib="DrosBAC"
/plasmid="pBelobAC11"
/note="end : SP6"

ORIGIN
Query Match          5.5%; Score 64; DB 9; Length 987;
Best Local Similarity 44.8%; Pred. No. 0.0026;
Matches 129; Conservative 24; Mismatches 135; Indels 0; Gaps 0;

QY 672 TTCTCTCGTTTTTTTTTAAATTTTTCATAAGTCTTTTGTATCTTCAATACAAATTTT 731
Db 761 TTTTWTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 820

QY 732 TCGCTGTATCTTGCAGAACTCTTCGATCATATCGCAATATACGTGAACACTGGTATCTA 791
```

```
Best Local Similarity 33.0%; Pred. No. 0.00073;
Matches 92; Conservative 69; Mismatches 118; Indels 0; Gaps 0;

QY 681 TTTTNTTAAATTTTTCATAAGTCTTTTCTTTTATCTTCAATACAAATTTTGGCTGTAT 740
Db 954 WWWWWWWWWTTTTTTTTTTTTTTTTTTTTTTTAAWMTTTTAAWMTTTTAAWMTTTT 895

QY 741 CTTGCAAACTCTTCGATCATATCGCAATATACGTGAACACTGGTATCTTAATTTGTGT 800
Db 894 TTTTWTATTTTWTATTTTWTATTTTWTATTTTWTATTTTWTATTTTWTATTTTWTAT 835

QY 801 GTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 860
Db 834 AAMTNTAAMTNTAAMTNTAAMTNTAAMTNTAAMTNTAAMTNTAAMTNTAAMTNTAAM 775

QY 861 TAAACATTTGTAAGTAAAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 920
Db 774 TTTAAATTTAATAATTTTCTTTAAAWATTTTAAAWATTTTAAAWATTTTAAAWATTT 715

QY 921 CAAAAATGAGATAGATACATTTTGAATTTTGTCTATTTT 959
Db 714 AAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 676

RESULT 7
CNS017V2          1101 bp DNA linear GSS 26-JUL-1999
LOCUS              Drosophila melanogaster genome survey sequence SP6 end of BAC
DEFINITION         BACN37L10 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION          AL108536
VERSION            AL108536.1 GI:5628840
KEYWORDS           Drosophila melanogaster (fruit fly)
SOURCE             Drosophila melanogaster
ORGANISM           Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE          1 (bases 1 to 1101)
AUTHORS            Genoscope.
TITLE              Direct Submission
JOURNAL            Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr)
COMMENT            - Web : www.genoscope.cns.fr
Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk - This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelobAC11.

FEATURES
source            1..1101
Location/Qualifiers
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACN37L10"
/clone_lib="DrosBAC"
/plasmid="pBelobAC11"
/note="end : SP6"

ORIGIN
Query Match          5.3%; Score 61.6; DB 9; Length 1101;
Best Local Similarity 44.8%; Pred. No. 0.0026;
Matches 129; Conservative 24; Mismatches 135; Indels 0; Gaps 0;

QY 672 TTCTCTCGTTTTTTTTTAAATTTTTCATAAGTCTTTTGTATCTTCAATACAAATTTT 731
Db 761 TTTTWTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 820

QY 732 TCGCTGTATCTTGCAGAACTCTTCGATCATATCGCAATATACGTGAACACTGGTATCTA 791
```


[illegible]

RESULT 12	CNS0243U	324 bp	DNA	linear	GSS 01-SEP-2000
LOCUS	CNS0243U/c				
DEFINITION	Tetraodon nigroviridis genome survey, sequence 17 end of clone 234115 of library G from Tetraodon nigroviridis, genomic survey sequence.				
ACCESSION	AL180291				
VERSION	AL180291.1	GI:7818348			
KEYWORDS	GSS; genome survey sequence.				
SOURCE	Tetraodon nigroviridis				
ORGANISM	Tetraodon nigroviridis				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontoidea; Tetraodontidae; Tetraodon.				
REFERENCE	1				
AUTHORS	Roest Crolius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fzames,C., Wincker,P., Brottier,P., Quetier,F., Saurin,W. and Weissenbach,J.				
TITLE	Estimate of human gene number provided by genome-wide analysis using Tetraodon nigroviridis DNA sequence				
JOURNAL	Nat. Genet. 25 (2), 235-238 (2000)				
MEDLINE	20296633				
PUBMED	10835645				

AUTHORS	Rozet Crollius, H., Jaillon, O., Dasilva, C., Ozouf-Costaz, C., Fimazes, C., Fischer, C., Bouneau, L., Billault, A., Quétier, F., Saurin, W., Bernot, A. and Weissenbach, J.
TITLE	Characterization and repeat analysis of the compact genome of the freshwater pufferfish tetraodon nigroviridis
JOURNAL	Genome Res. 10 (7), 939-949 (2000)
MEDLINE	20359837
PUBMED	10899143
REFERENCE	3 (bases 1 to 324)
AUTHORS	Genoscope.
TITLE	Direct Submission
JOURNAL	Submitted (12-APR-2000) Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
COMMENT	- Web : www.genoscope.cns.fr This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/tetraodon.

```

FEATURES             source
Location/Qualifiers
1..324
/organism="Tetraodon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"
/clone="234115"
/clone_lib="G"
/note="Genoscope sequence ID : C0AG234AE08LP1-end : T7"

```

Query Match 5.1%; Score 59.4; DB 9; Length 324;

	Best Local Similarity	46.2%;	Pred. No. 0.0073;	Matches 108;	Conservative 14;	Mismatches 112;	Indels 0;	Gaps 0;
Qy	660	TGTCACCTTGTTTCCCTCGTTTTTTTTTAATTTTCATAAGTCTCTTTGGTTTATCTTC	719					
Db	263	TTTCCTATTATTT	204					
Qy	720	AATACAAATTTTGGCTGTATCTTGCAAACTCTTCGATCATATCGCCAATATACGTGAAC	779					
Db	203	TTTTTATTTTTTTTGGTTTTTTTTTTTATAWTTTYYYYTTTTTNCONNNNTNCGNTTNN	144					
Qy	780	ACTGGTGATCTAATTTCTGCGTGAATGTGTAATTTAGATCTCAATCTCCGGTTTAAAA	839					
Db	143	NTTGGTGGGGTATATATGTTATTTGTTTTTACTGNNTNTTTTTTTTTTTTTTTTTTTTT	84					
Qy	840	GTGAATATATGATCATGTTAAAAACATTTGTAAGTAAGATGATAATAAAATGA	893					
Db	83	TTGAWTTTTTTTTTTTAAAGTGTAATKTWAAAANWAAANDBBKADNNNNNNNDRD	10	:	:	:	:	:

RESULT 13	CL648142	1187 bp	DNA	linear	GSS 06-JUL-2004
LOCUS	CL648142				
DEFINITION	CH213-154G18.SP6	CH213	<i>Gasterosteus aculeatus</i>	genomic clone	
	CH213-154G18	3',	genomic survey sequence.		
ACCESSION	CL648142				
VERSION	CL648142.1	GI:49667566			
KEYWORDS	GSS.				
SOURCE	<i>Gasterosteus aculeatus</i>	(three spined stickleback)			
ORGANISM	<i>Gasterosteus aculeatus</i>				

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1	Kingsley, D., Grimwood, J., Dickson, M., Schmutz, J. and Myers, R.M.	Expressed sequence tags from <i>Gasterosteus aculeatus</i>		
	1 (bases 1 to 1187)	Unpublished (2004)		
	Contact: Grimwood, Jane			
	Stanford Human Genome Center			
	Stanford University School of Medicine			
	975 S California Avenue, Palo Alto, CA 94304, USA			
	Tel: 650 320 5917			
	Fax: 650 320 5801			
	Email: jane@shgc.stanford.edu			
	Plate: 164			
	Class: BAC ends			
	High quality sequence start: 12			
	High quality sequence stop: 664.			

```

1.1187      /organism="Gasterosteus aculeatus"
            /mol_type="Genomic DNA"
            /strain="Salmon River"
            /db_xref="taxon:69293"
            /clone="CH213-164G18"
            /sex="Mixed"
            /cell_type="Blood"
            /clone_lib="CH213"
            /note="Vector: pRARBAC2.1; Site 1: EcoRI; The sequence of
the clone was established as a mapping and sequencing
collaboration at the Stanford Genome Evolution Center,
funded by the NIH Centers of Excellence in Genomic Science
(CEGS) initiative (http://cegs.stanford.edu). The clone
was isolated from the BAC library CHORI-213 built by
Pietter deJong in collaboration with the Stanford Genome
Evolution Center (http://www.chori.org/bacpac/). Clones
may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/ordering\_information.html)."
```

ORIGIN

Query Match 5.1%; Score 59.4; DB 9; Length 1187;
Best Local Similarity 51.4%; Pred. No. 0.0082;

Search completed: October 9, 2004, 08:37:32
Job time : 5783 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 8, 2004, 17:32:55 ; Search time 866 Seconds

(without alignments)
7019.434 Million cell updates/sec

Title: US-09-998-059-1

Perfect score: 1158

Sequence: 1 cacaaacatacactcaaat.....ctctcacaacaacaaaag 1158

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_23Sep04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001s:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002s:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003s:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1158	100.0	1164	10	ADJ19001 Thale cre
2	1158	100.0	1223	6	Abz17132 Arabidops
C 3	81	7.0	1118	3	Aac33222 Arabidops
C 4	65.4	5.6	10326	6	Abz134166 Human imm
C 5	56.2	4.9	6000	6	Abz133133 Human imm
C 6	56.2	4.9	40324	6	Abq67149 Human ang
7	55.8	4.8	5845	6	Abz133662 Human imm
8	54.6	4.7	7441	6	Abk40058 Human che
C 9	54.6	4.7	7479	6	Aas63345 Chemical
C 10	54.2	4.7	6713	6	Abz133080 Human imm
C 11	54.2	4.7	6713	6	Abz134526 Human met
C 12	54.2	4.7	6713	6	Abz170251 Chemical
C 13	53.6	4.6	7215	6	Abn91164 Staphyloc
C 14	53.4	4.6	29593	10	Abz37662 Human che
15	52.4	4.5	2000	8	Ada71938 Rice gene
16	52.2	4.5	337	8	Abx41811 Bovine ES
17	52	4.5	694	4	Aah53104 S. epider
18	52	4.5	875	4	Aai95044 Human neu
19	52	4.5	3028	4	Aah54496 S. epider
20	52	4.5	3032	4	Aah54863 S. epider
21	52	4.5	6968	4	Aah52683 S. epider

22	51.2	4.4	96588	9	ADA03026	AdA03026 Human MBN
23	51.2	4.4	96588	10	ADB72764	AdB72764 Human MBN
24	51.2	4.4	96588	10	ADC85506	AdC85506 Human Mbn
25	51.2	4.4	96588	12	ADM74621	AdM74621 Human car
C 26	51	4.4	11735	4	AAS45328	Aas45328 Chemical
C 27	51	4.4	11735	6	ABK28167	Abk28167 DNA trans
C 28	51	4.4	11735	6	AAS61141	Aas61141 Human gen
C 29	51	4.4	13511	6	ABL32280	AbL32280 Human imm
C 30	50.8	4.4	6621	8	ABZ10129	Abz10129 Haematopo
C 31	50.8	4.4	6621	8	ABZ10243	Abz10243 Haematopo
C 32	50.8	4.4	6621	10	ADB84153	AdB84153 Human lym
C 33	50.6	4.4	3683	8	ABZ10199	Abz10199 Haematopo
C 34	50.6	4.4	8011	6	ABL32078	AbL32078 Human imm
35	50.6	4.4	8011	6	AAD28368	Aad28368 Human che
36	50.4	4.4	419	8	ABX46069	Abx46069 Bovine ES
37	50.4	4.4	7544	4	AAS45300	Aas45300 Chemical
38	50.4	4.4	7544	6	ABK28139	Abk28139 DNA trans
39	50.2	4.3	7037	6	ABL33751	AbL33751 Human imm
40	50.2	4.3	13503	6	ABL34070	AbL34070 Human imm
41	50	4.3	597	6	ABQ36487	Abq36487 Oligonuc
C 42	50	4.3	597	6	ABQ36486	Abq36486 Oligonuc
43	50	4.3	5241	6	ABL70492	AbL70492 Chemical
44	50	4.3	5241	6	AAS61450	Aas61450 Human gen
C 45	49.8	4.3	556	5	ABV40163	Abv40163 Human pro

ALIGNMENTS

RESULT 1
ADJ19001
ID ADJ19001 standard; DNA; 1164 BP.
XX
AC ADJ19001;
XX
DT 20-MAY-2004 (first entry)
XX
DE Thale cress seed-specific promoter region Pl DNA.
XX
KW plant; seed-specific promoter region; seed development; food;
KW agricultural; thale cress; ds; Pl.
XX
OS Arabidopsis thaliana.
XX
PN US2003005485-A1.
XX
PD 02-JAN-2003.
XX
PF 30-NOV-2001; 2001US-00998059.
XX
PR 01-DEC-2000; 2000US-0250401P.
XX
(OHLR/) OHLROGGE J B.
PA (BENN/) BENNING C.
PA (GAOH/) GAO H.
PA (GIRK/) GIRKE T A A.
PA (WHIT/) WHITE J A.
XX
PI Ohlrogge JB, Benning C, Gao H, Girke TAA, White JA;
XX WPI; 2003-370848/35.
XX
DR New DNA comprising plant seed specific promoters, useful for controlling
XX expression of genes in plants, particularly for modifying seed products
XX (proteins, carbohydrates or oils), which are of major economic and food
XX values.
XX
PS Claim 1; SEQ ID NO 1; 57pp; English.
XX
CC The invention relates to a novel isolated DNA molecule which comprises a
CC plant seed-specific promoter region. The DNA molecule of the invention
CC may be useful for controlling the expression of genes in plants during
CC the different phases of seed development, in particular, for modifying

QY 1 CACAAACATACACTCAAAATCCAGACTCACATCTACTCAATTTATGCAACTTTCATCATGAA 60
Db |||||
QY 63 CACAAACATACACTCAAAATCCAGACTCACATCTACTCAATTTATGCAACTTTCATCATGAA 122
Db |||||
QY 61 AACATCAAAAACAGTCAGTAAGTAACAAAATCAAGTCAGATTCAGCACACAAAGCCAGTAA 120
Db |||||
QY 123 AACATCAAAAACAGTCAGTAAGTAACAAAATCAAGTCAGATTCAGCACACAAAGCCAGTAA 182
Db |||||
QY 121 GATAGAAAATTTAAGCAAGCTCATGCTAAGCTCGGCAAAATCTCTCTAATCAAAACAG 180
Db |||||
QY 183 GATAGAAAATTTAAGCAAGCTCATGCTAAGCTCGGCAAAATCTCTCTAATCAAAACAG 242
Db |||||
QY 181 TAACAACGAGTAATTAGCAAAATCCGAGCAGAAATCTCTACCCACCTCCGAAATTCACG 240
Db |||||
QY 243 TAACAACGAGTAATTAGCAAAATCCGAGCAGAAATCTCTACCCACCTCCGAAATTCACG 302
Db |||||
QY 241 TCTTCACTAAAATTTTGAAGGAATCGATCAATACCAACCCATTACACAAAATACATAA 300
Db |||||
QY 303 TCTTCACTAAAATTTTGAAGGAATCGATCAATACCAACCCATTACACAAAATACATAA 362
Db |||||
QY 301 TCAAAATGGCGAGAAATCGTACCTGAAACTTTGCTTCAAGTCGAGAGAGGAAAGGA 360
Db |||||
QY 363 TCAAAATGGCGAGAAATCGTACCTGAAACTTTGCTTCAAGTCGAGAGAGGAAAGGA 422
Db |||||
QY 361 AGATCGTGGAGAAAGGGGTTTAGGGTTTAAAGCTCAGACTTCTATTGGAGTAAATGGGACG 420
Db |||||
QY 423 AGATCGTGGAGAAAGGGGTTTAGGGTTTAAAGCTCAGACTTCTATTGGAGTAAATGGGACG 482
Db |||||
QY 421 GTGTCACTTTCCGTTTGGAAATGAACCTTTGGGCTCACGTTATGGGCTATTAGATATT 480
Db |||||
QY 483 GTGTCACTTTCCGTTTGGAAATGAACCTTTGGGCTCACGTTATGGGCTATTAGATATT 542
Db |||||
QY 481 TGATGGCTTTCTAGTAAATACAAATAGTTTATGGCTTAGTTTAAATAGCCCATGT 540
Db |||||
QY 543 TGATGGCTTTCTAGTAAATACAAATAGTTTATGGCTTAGTTTAAATAGCCCATGT 602
Db |||||
QY 541 TGGAAATATTGGACATGCTTTGGCTTACTAGTGTAAACATCAACCGACAGAGTTGTCG 600
Db |||||
QY 603 TGGAAATATTGGACATGCTTTGGCTTACTAGTGTAAACATCAACCGACAGAGTTGTCG 662
Db |||||
QY 601 AGCAAGTCGACGATATACAAATGGATCAAAACGCTAGTGTGCGCGCTCTCGCTCAT 660
Db |||||
QY 663 AGCAAGTCGACGATATACAAATGGATCAAAACGCTAGTGTGCGCGCTCTCGCTCAT 722
Db |||||
QY 661 GTGTCACTTTGTTCTCGTTTCTTAAATTTTCAATAGTTCTTTTGTATTATCTTCA 720
Db |||||
QY 723 GTGTCACTTTGTTCTCGTTTCTTAAATTTTCAATAGTTCTTTTGTATTATCTTCA 782
Db |||||
QY 721 ATACAAATTTTGGCTGTATCTTGCAAACTCTTCGATCATATCGCAATATAGTGAACA 780
Db |||||
QY 783 ATACAAATTTTGGCTGTATCTTGCAAACTCTTCGATCATATCGCAATATAGTGAACA 842
Db |||||
QY 781 CTGGTGATCTAATTTGTTGTTAATTTAGTAAATTTAGATTTCTCCGGTTTAAAG 840
Db |||||
QY 843 CTGGTGATCTAATTTGTTGTTAATTTAGTAAATTTAGATTTCTCCGGTTTAAAG 902
Db |||||
QY 841 TGAATTTATATGATCATGTTAAACATTTAGTAAAGTATGATATAATGATAAATTT 900
Db |||||
QY 903 TGAATTTATGATCATGTTAAACATTTAGTAAAGTATGATATAATGATAAATTT 962
Db |||||
QY 901 AGTTGATGATAACGTGAAGCAAAATAGATAGATATGATTTGTTGTTGTTGTTGTTG 960
Db |||||
QY 963 AGTTGATGATAACGTGAAGCAAAATAGATAGATATGATTTGTTGTTGTTGTTGTTG 1022
Db |||||
QY 961 ACATATGCGGAGAGTGTAGTACGCGCATGAAGATCAAGACACATTTGTCGAGCTCACAG 1020
Db |||||
QY 1023 ACATATGCGGAGAGTGTAGTACGCGCATGAAGATCAAGACACATTTGTCGAGCTCACAG 1082
Db |||||
QY 1021 AGTGACGTGTAAGACTTAGACTAGTCCCATGCAAACTTAATCTTACGTGGCTCAA 1080
Db |||||
QY 1083 AGTGACGTGTAAGACTTAGACTAGTCCCATGCAAACTTAATCTTACGTGGCTCAA 1142
Db |||||

QY 1081 ACCACGAGCTCAGCTTGCATATATATAAATCTCTCTAAGTCCCGTTCTTCTTCACTCATCT 1140
Db |||||
QY 1143 ACCACGAGCTCAGCTTGCATATATATAAATCTCTCTAAGTCCCGTTCTTCTTCACTCATCT 1202
Db |||||
QY 1141 CTCACAAACAAACAAAAG 1158
Db |||||
QY 1203 CTCACAAACAAACAAAAG 1220
Db |||||
RESULT 3
AAC33222/c
ID AAC33222 standard; DNA; 1118 BP.
XX AC AAC33222;
XX DT 17-OCT-2000 (first entry)
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 2235.
XX Arabidopsis thaliana.
XX Hybridisation assay; genetic mapping; gene expression control;
XX protein identification; signal transduction pathway; metabolic pathway;
XX promoter; termination sequence; ss.
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-00301439.
XX PR 25-FEB-1999; 99US-01218259.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.

[illegible]

Query Match

Thu Oct 14 09:37:08 2004

us-09-998-059-1.rng

CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 40324 BP; 12279 A; 436 C; 8127 G; 19482 T; 0 U; 0 Other;
 SQ

Query Match 4.9%; Score 56.2; DB 6; Length 40324;
 Best Local Similarity 49.5%; Pred. No. 0.018; Indels 0; Gaps 0;
 Matches 145; Conservative 0; Mismatches 148; Indels 0; Gaps 0;

669 TTGTTTCCTCGTTTCTTAAATTTTTCATAAGTTCCTTTGTTTATCTTCAATACAAAT 728
 12605 TTTTTCCTCGTTTCTTAAATTTTTCATAAGTTCCTTTGTTTATCTTCAATACAAAT 12664

729 TTTTGGCTGTATCTTCGAAACCTCTTCGATCATATGCCCAATATACGTGACACACTGAT 788
 12665 TTTTTCCTCGTTTCTTAAATTTTTCATAAGTTCCTTTGTTTATCTTCAATACAAAT 12724

789 CTAAATTTGTTGTTAAATTTTAAATTTAGATTCATCTCCGGTTTAAAGTGAATAT 848
 12725 TTTTTCCTCGTTTCTTAAATTTTTCATAAGTTCCTTTGTTTATCTTCAATACAAAT 12784

849 ATGTATCATGTTTAAACCAATTTGTAAGATGATAATAAATGATAAATTTAGTTGATG 908
 12785 GTTTTATTTAGTTTATTTTATAGTTTATTAATGAATTAATGAATTAATGAATTA 12844

909 GATAACGTGAAGCAAAAATGAGATGATACATTTGATTTTTCGTTATTTGA 961
 12845 ATAAATGGGTAGTATATAGTATTATGAAATATTTTATTATTGAGTTTATTA 12897

RESULT 7
 ABL33662
 ID ABL33662 standard; DNA; 5945 BP.
 XX AC ABL33662;
 XX DT 26-MAR-2002 (first entry)
 XX DE Human immune system associated gene SEQ ID NO: 1635.
 XX KW Human; immune system disease; cytosine methylation; antiasthmatic;
 XX KW antiarteriosclerotic; antianemic; cytostatic; neotropic;
 XX KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
 XX KW antirheumatic; antiarthritic; antidiabetic; antipsoriasis;
 XX KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
 XX KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
 XX KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
 XX KW ds.
 XX OS Homo sapiens.
 XX PN WO200200928-A2.
 XX PD 03-JAN-2002.
 XX PF 02-JUL-2001; 2001WO-EP007537.
 XX PR 30-JUN-2000; 2000DE-01032529.
 XX PR 01-SEP-2000; 2000DE-01043826.
 XX PA (EPIG-) EPIGENOMICS AG.
 XX PI Olek A, Piepenbrock C, Berlin K;
 XX DR WPI; 2002-130909/17.
 XX CC Nucleic acid comprising fragment of chemically modified gene, useful for
 PT diagnosis and treatment of diseases associated with abnormal cytosine
 PT methylation.
 XX PS Claim 1; SEQ ID NO 1635; 32pp + Sequence Listing; German.
 XX CC The present invention provides a number of human immune system associated
 CC genes which are modified by the methylation of cytosines. The sequences

Best Local Similarity 51.0%; Pred. No. 0.012;
 Matches 133; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

40 ATTATGCAACTTCATCATGAAACATCAAAACAGTCAAGTAACAAATCAAGTCAGAT 99
 4828 ATTAAAAACCCAAACCGTAAACCCCTAAACGACTTCAACGTTATTTAAACAACTCAAC 4769

100 TCAGCACACAAAGCCAGTAAAGATAGAAAATTTAACGAACGGTCTATGCTGAAGTCGGAA 159
 4768 TCAACCTCAAAACCCCTACAAAACAACTCTCTAAATTTATCTCAACTTCAACTCCCTCA 4709

160 AATATCTTCTTAATCAAAAACAGTAACAAACGAGTAATTTAGCAAAATCCGACGAAACTCT 219
 4708 ATAACTCACTAAATAAATCCAAACCACTATATAAAACCAAAACCGTAACTTAATCAAC 4649

220 CACCCACCTCCGAAATTCAGCTTCTCACTAAATTTTCGAAAGGAATCGATCAATACCAA 279
 4648 TAAAAACAAACCAAAACCGGTATCATTTCTAAATTTTAAACAACTACAAACAAACCAAA 4589

280 CCCATTACACAAAATACATAA 300
 4588 AACATCTACAAAATATAA 4568

RESULT 6
 ABQ67149
 ID ABQ67149 standard; DNA; 40324 BP.
 XX AC ABQ67149;
 XX DT 28-AUG-2002 (first entry)
 XX DE Human angiogenesis associated polynucleotide SEQ ID NO 179.
 XX KW Human; angiogenesis; methylation; eye disease; glaucoma; tumour;
 XX KW inflammation; rheumatoid arthritis; diabetic retinopathy; antiulcerds;
 XX KW macular degeneration; inflammatory bowel disease; Crohn's disease;
 XX KW antirheumatic; antiarthritic; antidiabetic; antipsoriasis;
 XX KW antiarteriosclerotic; ds.
 XX OS Homo sapiens.
 XX PN WO200246454-A2.
 XX PD 13-JUN-2002.
 XX PF 06-DEC-2001; 2001WO-EP014320.
 XX PR 06-DEC-2000; 2000DE-01061338.
 XX PA (EPIG-) EPIGENOMICS AG.
 XX PI Schacht O;
 XX DR WPI; 2002-500450/53.
 XX PT New nucleic acid fragments from chemically treated angiogenesis-
 PT associated genes, useful for determining methylation status, e.g. in
 PT diagnosis or treatment of cancer.
 XX PS Claim 1; SEQ ID NO 179; 41pp + Sequence Listing; German.
 XX CC The invention relates to a nucleic acid (I) comprising a segment of 18
 CC bases of chemically pretreated DNA of angiogenesis-associated genes (II)
 CC having sequences (ABQ66971-ABQ67178) or their complements. (I), also
 CC related oligomers, are used to evaluate the methylation status and/or
 CC single-nucleotide polymorphisms, in angiogenesis-related genes, for
 CC diagnosis and treatment of eye diseases, proliferative retinopathy,
 CC neovascular glaucoma, solid tumours, inflammation, rheumatoid arthritis,
 CC diabetic retinopathy, macular degeneration caused by neovascularisation,
 CC psoriasis, arteriosclerosis, inflammatory bowel diseases, ulcers and
 CC Crohn's disease. Note: The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format

XX	PD	18-OCT-2001.
XX	PP	06-APR-2001; 2001WO-EP004016.
XX	PR	06-APR-2000; 2000DE-01019058.
XX	PT	07-APR-2000; 2000DE-01019173.
XX	PP	30-JUN-2000; 2000DE-01032529.
XX	PR	01-SEP-2000; 2000DE-01043826.
XX	PP	(EPIG-) EPIGENOMICS AG.
XX	PP	Olek A, Piepenbrock C, Berlin K;
XX	DR	WPI; 2002-010834/01.
XX	PP	New nucleic acid, useful for diagnosis and therapy of metabolic disease,
XX	PP	solid tumor and cancers, comprises segment of chemically modified genomic
XX	PP	sequences of genes associated with metabolism.
XX	PS	Claim 1; Page 109-111; 143pp; English.
XX	PS	The invention relates to a nucleic acid (I) comprising a sequence at
XX	CC	least 18 bases of a segment of the chemically pretreated DNA of genes
XX	CC	associated with metabolism such as DUSP2 (NM_004418), EPHX2 (NM_001979),
XX	CC	QDPR (NM_000320), SGSH (NM_000199), SHMT2 (NM_005412), SLC7A2
XX	CC	(NM_003046), SLC7A4 (NM_004173) and TYMS (NM_001071) (all undefined). (I)
XX	CC	are useful for diagnosis and therapy of metabolic disease, solid tumours
XX	CC	and cancers; as primer oligonucleotides for the amplification of DNA
XX	CC	sequences; for detecting the cytosine methylation state and/or single
XX	CC	nucleotide polymorphisms (SNPs) in a chemically treated DNA of genes
XX	CC	associated with metabolism. An array of (I) is useful for ascertaining
XX	CC	genetic and/or epigenetic parameters for the diagnosis and/or therapy of
XX	CC	existing diseases or the predisposition to specific diseases by analysing
XX	CC	cytosine methylations. The method involves chemically treating genomic
XX	CC	DNA sample by a solution of bisulphite, hydrogen sulphite or disulphite
XX	CC	such that cytosine bases which are unmethylated at the 5th-position are
XX	CC	converted to uracil or another base which is dissimilar to cytosine in
XX	CC	terms of hybridisation behaviour and amplifying fragments of the
XX	CC	chemically pretreated genomic DNA. The genomic DNA is from cells or
XX	CC	cellular components which contain DNA, sources of DNA comprising, for
XX	CC	e.g. cell lines, biopsies, blood, sputum, stool, urine, cerebral-spinal
XX	CC	fluid, tissue embedded in paraffin such as tissue from eye, intestine,
XX	CC	kidney, brain, heart, prostate, lung, breast or liver, histologic object
XX	CC	slides and their combinations. Genetic parameters are mutations, in
XX	CC	particular insertions, deletions, point mutations, inversions and
XX	CC	polymorphisms of genes associated with metabolism and sequences further
XX	CC	required for their regulation. Epigenetic parameters are in particular
XX	CC	cytosine methylations and further chemical modifications of DNA bases of
XX	CC	genes associated with metabolism. Further epigenetic parameters include
XX	CC	for e.g. the acetylation of histones which correlates with DNA
XX	CC	methylation. AA563306-AA563373 represent chemically pretreated metabolism
XX	CC	associated genes, and related primers of the invention
XX	SS	Sequence 7479 BP; 2133 A; 68 C; 1273 G; 4005 T; 0 U; 0 Other;
XX	SS	Query Match 4.7%; Score 54.6; DB 6; Length 7479;
XX	SS	Best Local Similarity 48.8%; Pred. No. 0.028;
XX	SS	Matches 147; Conservative 0; Mismatches 154; Indels 0; Gaps 0;
QY	DB	657 TCATGTGCACCTTGTTTCCTCGTTTTTTTTTAAATTTCATAAGTCTCTTTGGTTTATC 716
Db	DB	6941 TTATTTATTTTGTGTTTAAATTAATCATGTTTTTTTTTATATATTTTGA 7000
QY	DB	717 TTCATACAAATTTTGGCTGTAATCTGCAACTTCGATCATATCGCAATACGNG 776
Db	DB	7001 TTTTAAAAATTTTGTGTTTATTTTCGGAAAATTCGATATTTTATTTTAGTTAGTT 7060
QY	DB	777 AACACTGGTGATCTAAATTTGTTGTTAAATTTGTTAAATTTAGATTTCTATCTCCGTTTA 836
Db	DB	7061 TATTTGGCGGTATATTTTAAAGGTTTATAGATAGTTTTTTTATTTTTCGATAGGTTTT 7120
QY	DB	837 AAAGTGAATATATGATCATGTTTAAACAATTCGTAAGTAGATGATATAAATGNTAA 896

Db	7121	TGTTAGTAATTTTGTGTTATTTTATTTTAAATAATTTTATTTTAAATATGTTAATAA 7181
QY	897	ATTTAGTTGATGATAACGTCGAACAAAAATAGATAGATACATTTGATTTTGTGCTAT 956
Db	7181	TTTGTGTTTGTGTTAGTTGTTTAAAGAGATTGTAATATTTATATATATTTTATTTT 7240
QY	957	T 957
Db	7241	T 7241
RESULT 10		
ABL33080/c	ABL33080 standard; DNA; 6713 BP.	
ID		
XX	ABL33080;	
AC	26-MAR-2002 (first entry)	
XX	Human immune system associated gene SEQ ID NO: 1053.	
DT		
XX	Human; immune system disease; cytosine methylation; antiasthmatic;	
DE	antiartherosclerotic; antianaemia; cytostatic; nootropic;	
XX	neuroprotective; anti-HIV; anticonvulsant; ophthalmological;	
KW	antiheumatic; antiarthritic; antidiabetic; antipsoriatic;	
KW	antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;	
KW	acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;	
KW	neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;	
ds.		
XX	Homo sapiens.	
OS	WO200200928-A2.	
XX	03-JAN-2002.	
PN	02-JUL-2001; 2001WO-EP007537.	
PD	30-JUN-2000; 2000DE-01032529.	
PF	01-SEP-2000; 2000DE-01043826.	
XX	(EPIG-) EPIGENOMICS AG.	
XX	Olek A, Piepenbrock C, Berlin K;	
PI	WPI; 2002-130909/17.	
XX	Nucleic acid comprising fragment of chemically modified gene, useful for	
DR	diagnosis and treatment of diseases associated with abnormal cytosine	
XX	methylation.	
PT	Claim 1; SEQ ID NO 1053; 32pp + Sequence Listing; German.	
PT	The present invention provides a number of human immune system associated	
XX	genes which are modified by the methylation of cytosines. The sequences	
CC	can be used in the diagnosis and treatment of immune system disorders,	
CC	including eye diseases such as retinopathy, neovascular glaucoma and	
CC	macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid	
CC	leukemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,	

Thu Oct 14 09:37:08 2004

Db 720 MTGMYSGRYWTSWKYKCKSWKYSRMWYWSWMAKTWKMRRYATRMWYRYSMKW 779
QY 830 CGGTTTAAAGTGAATTATATGATCATCGTTAAACATTGTAAGTAAGATGATAATAA 889
Db 780 YTWCTWGYWYWNRYTKGRYMWYKCTKYWYWSATYWTGTWAAWMAKTWRMGMTGA 839
QY 890 ATGATAAATTTAGTTGATGATAACGTGAAGCAAAAAATGAGATAGATACATTGATTT 949
Db 840 KTRGRARKARYWVKWATWCATKRWMTKGKAKWATWMAKWKYYSWNRWYKYKTR 899
QY 950 GTGCTATTTGACATATGCGGAGAGTCAGCTACGCCATGAAGATCAAGAGACACTTGCT 1009
Db 900 RT-RYKTCWKAWSWAYWRWYWKSGAKWMMWKGGRWGTKYWYCTTWKACGRAT 958
QY 1010 CGAGCTCACAGAGTGTGCTGTAAAGCTTAGACTGAA 1047
Db 959 KYMCCAGWAWMYSYSWTRTYWRTWYRWWASRTAKEM 996

Search completed: October 9, 2004, 04:43:59
Job time : 870 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 9, 2004, 03:54:36 ; Search time 145 Seconds
(without alignments)
5676.507 Million cell updates/sec

Title: US-09-998-059-1
Perfect score: 1158
Sequence: 1 cacaacatacactcaaat.....ctctcacaacacacaaaag 1158

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/prodata/1/ina/5A COMB.seq.*
- 2: /cgn2_6/prodata/1/ina/5B COMB.seq.*
- 3: /cgn2_6/prodata/1/ina/6A COMB.seq.*
- 4: /cgn2_6/prodata/1/ina/6B COMB.seq.*
- 5: /cgn2_6/prodata/1/ina/PCUS COMB.seq.*
- 6: /cgn2_6/prodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	53.6	4.6	7215	3	US-09-134-001C-627
2	52	4.5	694	4	US-09-710-279-1601
3	52	4.5	3028	4	US-09-710-279-3860
4	52	4.5	3032	4	US-09-710-279-4227
5	52	4.5	6968	4	US-09-710-279-759
6	49.8	4.3	4140	3	US-08-894-731-2
7	48.8	4.2	1851	4	US-09-601-198-51
8	48.6	4.2	1141	4	US-09-806-708B-22
9	47.4	4.1	1141	4	US-09-806-708B-22
10	45.8	4.0	7218	1	US-08-232-463-14
11	45.6	3.9	8607	4	US-10-204-708-71
12	45.4	3.9	11049	4	US-10-204-708-23
13	44.4	3.8	6801	4	US-10-204-708-62
14	43.6	3.8	3600	3	US-08-894-731-1
15	43.2	3.7	1316	4	US-09-270-767-2481
16	43.2	3.7	1316	4	US-10-204-708-59
17	43	3.7	640681	4	US-09-790-988-1
18	42.8	3.7	6156	4	US-10-204-708-77
19	42.4	3.7	6020	4	US-09-543-398B-11
20	42.2	3.6	5915	4	US-09-543-398B-9
21	42	3.6	66933	4	US-09-543-398B-11
22	42	3.6	66933	4	US-09-543-398B-9
23	42	3.6	72049	4	US-09-543-398B-9
24	42	3.6	72049	4	US-09-543-398B-9
25	41.8	3.6	519	4	US-09-248-796A-6405
26	41.4	3.6	658	3	US-08-998-416-595
27	41.4	3.6	51952	3	US-08-947-823-1

C 28	41.2	3.6	10144	4	US-10-204-708-94	Sequence 94, Appl
C 29	41.2	3.6	10467	4	US-10-204-708-2	Sequence 2, Appl
C 30	41	3.5	8607	4	US-10-204-708-72	Sequence 72, Appl
C 31	40.6	3.5	5562	4	US-10-204-708-63	Sequence 63, Appl
C 32	40.6	3.5	6156	4	US-10-204-708-60	Sequence 60, Appl
C 33	40.6	3.5	7218	1	US-08-232-463-14	Sequence 14, Appl
C 34	40.6	3.5	7786	4	US-09-790-988-2	Sequence 2, Appl
C 35	40.6	3.5	8607	4	US-10-204-708-72	Sequence 72, Appl
C 36	40.6	3.5	11049	4	US-10-204-708-23	Sequence 23, Appl
C 37	40.4	3.5	600	4	US-09-601-198-142	Sequence 142, App
C 38	40.4	3.5	1103	3	US-09-122-400B-20	Sequence 20, Appl
C 39	40.4	3.5	1167	1	US-08-592-658-1	Sequence 1, Appl
C 40	40.4	3.5	2672	1	US-08-703-947-1	Sequence 1, Appl
C 41	40.4	3.5	6040	4	US-10-204-708-69	Sequence 69, Appl
C 42	40.2	3.5	6306	4	US-10-204-708-50	Sequence 50, Appl
C 43	40.2	3.5	640681	4	US-09-790-988-1	Sequence 1, Appl
C 44	40.2	3.5	786431	4	US-09-751-389-3	Sequence 3, Appl
C 45	40	3.5	4832	4	US-09-457-037B-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1

US-09-134-001C-627
; Sequence 627, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 627
; LENGTH: 7215
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-627

Query Match 4.6%; Score 53.6; DB 3; Length 7215;
Best Local Similarity 52.2%; Pred. No. 0.00061;
Matches 119; Conservative 0; Mismatches 109; Indels 0; Gaps 0;
QY 697 ATAGTCTCTTTGTTTATCTTCAATACAAATTTTGGCTGTATCTTGCAAACTCTTGA 756
DB 372 ATTACCTCTGTTAAATTTAAGATTATCAATTTCTGATGAAGCTTACCTACTTTTGA 431
QY 757 TCATATCCCAATATACGTGAACACTGGTGATCTAAATTTGTTGTTAAATTTTAAATTT 816
DB 432 TTTTCCGCTACTATTTTAAATGAAGTCAATTAATCAATTTTACAACTTAATAT 491
QY 817 AGATTCTATTCTCCGGTTTAAAGTGAATTATATATATATATATATATATATATATAT 924
DB 492 CGCTATACCACTCTTTTAAAGTGAATATAGTATCTCGGATTTTATATATATATAT 551
QY 877 AGATGATTAATAATGATTAATTTAGTTTGTGATGATTAACGTGAAGCAAA 924
DB 552 AGAAATGAATCAAAAGATGATCAAAATCAAGTTGTGTGTCATCAAA 599

RESULT 2

US-09-710-279-1601
; Sequence 1601, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS

FILE REFERENCE: PU3480US
CURRENT APPLICATION NUMBER: US/09/710,279
CURRENT FILING DATE: 2000-11-09
PRIOR APPLICATION NUMBER: 60/164,258
PRIOR FILING DATE: 1999-11-09
NUMBER OF SEQ ID NOS: 4472
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1601
LENGTH: 694
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
OTHER INFORMATION: nucleic acid sequence
US-09-710-279-1601

Query Match 4.5%; Score 52; DB 4; Length 694;
Best Local Similarity 51.8%; Pred. No. 0.00068;
Matches 118; Conservative 0; Mismatches 110; Indels 0; Gaps 0;

QY 697 ATAAGTCTCTTTGTTTATCTTCAATACAAATTTTGGGTGATCTTGCACAACTCTCGA 756
Db 246 ATTACCTCTGTTTAAATTTAAGATATATCAATTTCTAGATGGCGCTACCTACTTTTGA 305

QY 757 TCATATCGCCAAATATACGTGAACACTGGTGATCTAATTTGTTGTTAATTTTAAATTT 816
Db 306 TTTCACGCTACTATTTTAAATGAAAGTCAATTAATCTCCATTTTACACAAATTAATAT 365

QY 817 AGATTCATCTCCGGTTTAAAGTGAATTTATATGATCATGTTAAACATTTGAAGTA 876
- Db 366 TGCTTATACCACTCTTTTAAAGAGTGAATATAGTATCTCGAATTTTATTAATTTGATTA 425

QY 877 AGATGATAATAAATGATAATTTAGTTGATGATAACGTGAAGCAA 924
Db 426 AGAATGAATCAAAAGATGATCAAAATCAAGTTGTGTCATCAA 473

RESULT 3
US-09-710-279-3860
Sequence 3860, Application US/09/710279
Patent No. 6703492
GENERAL INFORMATION:
APPLICANT: KIMMERLY, WILLIAM JOHN
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
FILE REFERENCE: PU3480US
CURRENT APPLICATION NUMBER: US/09/710,279
CURRENT FILING DATE: 2000-11-09
PRIOR APPLICATION NUMBER: 60/164,258
NUMBER OF SEQ ID NOS: 4472
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3860
LENGTH: 3028
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
OTHER INFORMATION: nucleic acid sequence
US-09-710-279-3860

Query Match 4.5%; Score 52; DB 4; Length 3028;
Best Local Similarity 51.8%; Pred. No. 0.0011;
Matches 118; Conservative 0; Mismatches 110; Indels 0; Gaps 0;

QY 697 ATAAGTCTCTTTGTTTATCTTCAATACAAATTTTGGGTGATCTTGCACAACTCTCGA 756
Db 2580 ATTACCTCTGTTTAAATTTAAGATATATCAATTTCTAGATGGCGCTACCTACTTTTGA 2639

QY 757 TCATATCGCCAAATATACGTGAACACTGGTGATCTAATTTGTTGTTAATTTTAAATTT 816
Db 2640 TTTCACGCTACTATTTTAAATGAAAGTCAATTAATCTCCATTTTACACAAATTAATAT 2699

QY 817 AGATTCATCTCCGGTTTAAAGTGAATTTAGTTGATGATAACGTGAAGTA 876

FILE REFERENCE: PU3480US
CURRENT APPLICATION NUMBER: US/09/710,279
CURRENT FILING DATE: 2000-11-09
PRIOR APPLICATION NUMBER: 60/164,258
NUMBER OF SEQ ID NOS: 4472
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1601
LENGTH: 694
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
OTHER INFORMATION: nucleic acid sequence
US-09-710-279-1601

Query Match 4.5%; Score 52; DB 4; Length 3032;
Best Local Similarity 51.8%; Pred. No. 0.0011;
Matches 118; Conservative 0; Mismatches 110; Indels 0; Gaps 0;

QY 697 ATAAGTCTCTTTGTTTATCTTCAATACAAATTTTGGGTGATCTTGCACAACTCTCGA 756
Db 2689 ATTACCTCTGTTTAAATTTAAGATATATCAATTTCTAGATGGCGCTACCTACTTTTGA 2748

QY 757 TCATATCGCCAAATATACGTGAACACTGGTGATCTAATTTGTTGTTAATTTTAAATTT 816
Db 2749 TTTCACGCTACTATTTTAAATGAAAGTCAATTAATCTCCATTTTACACAAATTAATAT 2808

QY 817 AGATTCATCTCCGGTTTAAAGTGAATTTAGTTGATGATAACGTGAAGTA 876
Db 2809 TGCTTATACCACTCTTTTAAAGTGAATATAGTATCTCGAATTTTATTAATTTGATTA 2868

QY 877 AGATGATAATAAATGATAATTTAGTTGATGATAACGTGAAGCAA 924
Db 2869 AGAATGAATCAAAAGATGATCAAAATCAAGTTGTGTCATCAA 2916

RESULT 5
US-09-710-279-759
Sequence 759, Application US/09/710279
Patent No. 6703492
GENERAL INFORMATION:
APPLICANT: KIMMERLY, WILLIAM JOHN
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
FILE REFERENCE: PU3480US
CURRENT APPLICATION NUMBER: US/09/710,279
CURRENT FILING DATE: 2000-11-09
PRIOR APPLICATION NUMBER: 60/164,258
NUMBER OF SEQ ID NOS: 4472
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 759
LENGTH: 6968
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
OTHER INFORMATION: nucleic acid sequence
US-09-710-279-759

FILE REFERENCE: PU3480US
CURRENT APPLICATION NUMBER: US/09/710,279
CURRENT FILING DATE: 2000-11-09
PRIOR APPLICATION NUMBER: 60/164,258
NUMBER OF SEQ ID NOS: 4472
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1601
LENGTH: 694
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
OTHER INFORMATION: nucleic acid sequence
US-09-710-279-1601

Query Match 4.5%; Score 52; DB 4; Length 3032;
Best Local Similarity 51.8%; Pred. No. 0.0011;
Matches 118; Conservative 0; Mismatches 110; Indels 0; Gaps 0;

QY 697 ATAAGTCTCTTTGTTTATCTTCAATACAAATTTTGGGTGATCTTGCACAACTCTCGA 756
Db 2689 ATTACCTCTGTTTAAATTTAAGATATATCAATTTCTAGATGGCGCTACCTACTTTTGA 2748

QY 757 TCATATCGCCAAATATACGTGAACACTGGTGATCTAATTTGTTGTTAATTTTAAATTT 816
Db 2749 TTTCACGCTACTATTTTAAATGAAAGTCAATTAATCTCCATTTTACACAAATTAATAT 2808

QY 817 AGATTCATCTCCGGTTTAAAGTGAATTTAGTTGATGATAACGTGAAGTA 876
Db 2809 TGCTTATACCACTCTTTTAAAGTGAATATAGTATCTCGAATTTTATTAATTTGATTA 2868

QY 877 AGATGATAATAAATGATAATTTAGTTGATGATAACGTGAAGCAA 924
Db 2869 AGAATGAATCAAAAGATGATCAAAATCAAGTTGTGTCATCAA 2916

	Matches	93;	Conservative	342;	Mismatches	476;	Indels	2;	Gaps	2;
QY	226	CCTCCGAAATTCACGTCTTCACTAAAATTTTGAAAGGAATCGATCAATAACCAOCCCAT	T	285	: : : : : : : : : : : : : : :					
Db	53	MSKSRKTWARMYCKYERWYNNSRWKKWGWYKKWYBCANNSTSBRYHARRWKDMKTAYBM	:	112	: : : : : : : : : : : : : : :					
QY	286	ACA CA A A A T A C A T A A T C A A A A T G C G A G A A T C G T A C T C G A A A C T T T G C T T C A A G T C G C A	T	345	: : : : : : : : : : : : : : :					
Db	113	TMTNKGKGTGWRHRMYEWRAMBTDVTHHYVTAMNNNAWTTCMMDKDDKRTAWWWKKNNNA	:	172	: : : : : : : : : : : : : : :					
QY	346	GAGAGAGAAAAAGGAAGTCTGGAGAAAAGGGTTTTAGGGTTTTAAGCTCAGACTCTTAATT	T	405	: : : : : : : : : : : : : : :					
Db	173	TGWD DDTKYHHMNNNGCBTVTWVRYKTRDWSBKRMNYGMWBWKNWSVDVIYYWVVWD	D	232	: : : : : : : : : : : : : : :					
QY	406	GGAGTAAATGGGACGGGTGCACATTTTCCGTTT - TGSAATAGAACCTTTGGGCTCACGTTA	A	464	: : : : : : : : : : : : : : :					
Db	233	MCKRKVRWVETRGRMRNMVMVAETAHRRVYNGWTBAMAYRRTWMNNNNNAKAMCKEA	A	292	: : : : : : : : : : : : : : :					
QY	465	TGGGCTATATAGATATTTGATGGGCTTTCTAGTAAATACAAATATAAGCTTATTCGGCTTAGT	T	524	: : : : : : : : : : : : : : :					
Db	293	KYWGWNRAVBASTCTTKWSKTTKVRTSCHANNCRAGDANKDHKWKSAAWGYVWNNNN	N	352	: : : : : : : : : : : : : : :					
QY	525	TTAAATAAGCCCATGTTTGGAAATATTTGACACATGCTTCGGCTACTAGTCTTAAACATCG	C	584	: : : : : : : : : : : : : : :					
Db	353	NNWTYKARHARDWDWHSSAKWKHANADAHYSKKWTBYKRKTWNNNGITTMKRMWA	A	412	: : : : : : : : : : : : : : :					
QY	585	AAACCGAACAGTTGTGAGACAAGTCGCAGCATATACAAATGATCAACACGCCCTAGTGTG	C	644	: : : : : : : : : : : : : : :					
Db	413	WYWKMDMBGTYNNNNNGRTYYGWTKKQWYYKWKANNCWKRAWDHKTCTHNNTTW	W	472	: : : : : : : : : : : : : : :					
- QY	645	GCGGCGTCTCGTCAATGTGACCTGTTCTCGTTTCTGTTTTTTTAAATTTTTCATAAGTTC	C	704	: : : : : : : : : : : : : : :					
Db	473	WKOKTYNNCYWKSMTNGKSHRBAAVYTWTWMMRRYAHANNNDWYMKWACTWXYBV	V	532	: : : : : : : : : : : : : : :					
- QY	705	TTTTGTTTATCTTCAATACAAATTTTTGGCTGTATCTTGCAAACTCTTCGATCATATCG	C	764	: : : : : : : : : : : : : : :					
Db	533	CSKNNNNYAATYKSGSNWTSRYHWKTNNSWRPSTDTRSMGRANNYABABHYGYKNWTR	R	592	: : : : : : : : : : : : : : :					
QY	765	CCAATATACGTGAACACCTGGTGATCAATATTTGTGTGTTAATGTTTAAATTTAGATTCTA	A	824	: : : : : : : : : : : : : : :					
Db	593	WBEWSHTWBHBRAAGAAHYWMBMWBAKCHCMKAWYKAKYAGAGGSNNNNNNNNNNNN	N	652	: : : : : : : : : : : : : : :					
QY	825	TTCTCCGGTTTAAAAGTGAATATATATGATCATGGTTAAAACATGTTAAGTAAGATGATA	A	884	: : : : : : : : : : : : : : :					
Db	653	NNATCARDYY - AASRWYAMANAKEYYYKBANNAIYTHANNWGCWNNNATDTRRTWKN	N	711	: : : : : : : : : : : : : : :					
QY	885	ATAAATGATAAATTTAGTTGATGATAACGTGAAGCAAAAAATCGAGATAGATACATTG	T	944	: : : : : : : : : : : : : : :					
Db	712	NNNNAGTWNNNNNAKVASAKNYAAAAVKAAKHWRWANKWAMRGHADAAABITDK	K	771	: : : : : : : : : : : : : : :					
QY	945	ATTTTCTCGTATTTGACATATCGCGAGAGTGAGCTACGCGCATGAAGATCAAGAGACAC	A	1004	: : : : : : : : : : : : : : :					
Db	772	RNNGAYTKVTTTTNNNTYRGVVNTAAARDGANNNNNNNNNNNNNNNNNNNNNNNNNNN	N	831	: : : : : : : : : : : : : : :					
QY	1005	TTGCTCGAGCTACAGAGTGCAGTGTAAAGACCTTAGACTGAAGTCCCCTGCAAACTA	A	1064	: : : : : : : : : : : : : : :					
Db	832	YGFTNNNNNNNNAYAWWTKWYTTDDRWRBRAYTNNNNNNRMYIGAYADDIYYMSDT	T	891	: : : : : : : : : : : : : : :					
QY	1065	ATCCTACGTGGCTCAAACCACGAGTCTCACTTGACAAATATATAAACTCCTCTTAAGTCCCG	C	1124	: : : : : : : : : : : : : : :					
Db	892	CDAMKWDATKMNNATYYNRGTAWRIINNNNNNWTKTYBYBHAAANNNNNNNGKCTAHTW	H	951	: : : : : : : : : : : : : : :					
QY	1125	TTCTCTTCATCCA		1137	: : : : : : : : : : : : : : :					
Db	952	WVCATKTKTGCW		964	: : : : : : : : : : : : : : :					

RESULT 9

301	VTYNWCWRMTYMGKT--MTNNNNNNKAWYRTKTVACNRRYYDTAVWIBGRNTRKICIA	Db
727	ATTTTGGCTGTATCTTGCACAACTCTTCGATCATATCGCCCATATACGTGAACACTGGTG	Qy
243	YBWIYEMYGKHHWBWRRABHSWNWVKCNKYVSVHMYHAMRYBKWBABAVGCNNNW	Db

[illegible]

RESULT 10
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Alexandria Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA

MIT: 22513-0299
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/232,463
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/07/935,313
 FILING DATE:
 APPLICATION NUMBER: EP 91 114 300.6
 FILING DATE: 26-AUG-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: BENT, Stephen A.
 REGISTRATION NUMBER: 29,768
 REFERENCE/DOCKET NUMBER: 30472/114 IMMU
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703)836-9300
 TELEFAX: (703)683-4109
 TELEX: 899149

```

/ INDEX: 895149
/ INFORMATION FOR SEQ ID NO: 14:
/ SEQUENCE CHARACTERISTICS:
/     LENGTH: 7218 base pairs
/     TYPE: nucleic acid
/     STRANDEDNESS: single
/     TOPOLOGY: linear
/     IMMEDIATE SOURCE:
/     CLONE: pTZ9pt-Fls
/ US-08-232-463-14

```

[illegible][illegible]

```

RESULT 11
US-10-204-708-71/c
; Sequence 71, Application US/10204708
; Patent No. 667731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; TITLE OF INVENTION: by Assessing DNA Methylation
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 71
; LENGTH: 8607
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-71

```

Query Match	3.9%;	Score 45.6;	DB 4;	Length 8607;
Best Local Similarity	49.0%;	Pred. No. 0.069;		
Matches 149;	Conservative 0;	Mismatches 154;	Indels 1;	Gaps 1;
QY	5	AACATACACTCAAAATCCAGAGCTCACATCTACTCAATTTATGCAACTTCATCATGAAAAACA	64	
DBb	5464	AAATTAATTACTTTTATCACAACTTAATTTTCATTCTCATAAAAATCTTTAAAAACCCCAAACT	5405	
QY	65	TCAAAAAACAGTCAAAAGTAAACAAAATCAAGTCAGATTTCAGCACACAAAAGCCAGTTAAAGATA	124	
DBb	5404	CGAAAAATAATAAAAAAATAAAACATAAAAAATAAAATAAAATAAAAAAATAAAAAA	5345	
QY	125	GAAATTTTAAACGACGCTCATGCTTAAGCTGCGCAAAATACTTCCTTAATCAAAACAGATTAAC	184	
DBb	5344	AACGATTAATAAAAAACAATAAACAACAAAAACAATAAAAAAATAAAAAAATAAAAAATA	5285	
QY	185	AACGAGTAATTTAGCAAAATTCGGAGC-AGAAAACTCTCACCCACCTCCGAAATTCACGTCT	243	
DBb	5284	AATAAATAATAAAAAAATAAAATCAACGAAACTCAAAATTCAAATTCCTTAAAAACCCCAAT	5225	
QY	244	TCACTAAATTTTCGAAGGGAATCGATCAATTAACCAACCCCAATTCACAAAATACATAATCA	303	

Thu Oct 14 09:37:08 2004

us-09-998-059-1.rni

Db 5224 CTCCAACTTTCTCATAAGCATCCTCGATCCGACCTCCTAACTAAACCCACCACCA 5165

QY 304 AAAT 307
Db 5164 TAAT 5161

RESULT 12
US-10-204-708-23
; Sequence 23, Application US/10204708
; Patent No. 667731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; TITLE OF INVENTION: by Assessing DNA Methylation
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 23
; LENGTH: 11049
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

US-10-204-708-23
Query Match 3.9%; Score 45.4; DB 4; Length 11049;
Best Local Similarity 49.5%; Pred. No. 0.084;
Matches 144; Conservative 0; Mismatches 146; Indels 1; Gaps 1;
QY 669 TTGTTCTCCTGTTTTTTAAATTTTTCATAGTCTTTGTTTAACTCAATCAAAAT 728
Db 7527 TTTTATTCTGAAATTTAGATTTATTAGGAATTTTTTTTTTTGTTAGAGTTTCGAGAAAGA 7586
QY 729 TTTGGCGTATCTTGCNAACCTCTCGATCATATCGCAATATACGTGACACACTGGTGAT 788
Db 7587 TGTGTTGTGTTTTTTGTTTATTTTGTTCGGAATTTTTTATTTGTTGTTTTTTTGGT 7646
QY 789 CTAAATTTGTTGTTAAATTTGTTAAATTTAGATTTCTCCGGTTTAAAGTGAATAT 848
Db 7647 ATATTTTTCGGTTATTTTTCGTTTTTTTGTATTTAGTTATACGGTTTGTAGTTTTTTT 7706
QY 849 ATGTATCATGGTTAAACATTTGTAAGTAGATGATAAATAAATGATAATTTAGTTGATG 908
Db 7707 GTTT-TAGTTTTTTTTTTTAAATTTTAAATAAATTTATTTATTTATTTATTTAGTTGTT 7765
QY 909 GATAACGTGAACAAATAATGATAGATACATTTGATTTCGCTATTTT 959
Db 7766 TTTTGTAGAGTAAGTTTTTTTAAAGGATAGGTATTTTGTGCGTTTTT 7816

RESULT 13
US-10-204-708-62
; Sequence 62, Application US/10204708
; Patent No. 667731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication

; TITLE OF INVENTION: by Assessing DNA Methylation
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 62
; LENGTH: 6801
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

US-10-204-708-62
Query Match 3.8%; Score 44.4; DB 4; Length 6801;
Best Local Similarity 48.8%; Pred. No. 0.13;
Matches 120; Conservative 0; Mismatches 126; Indels 0; Gaps 0;
QY 660 TGTCACCTGTTTCTCTCGTTTTTTTAAATTTTTCATAAGTTCTTTTGTATCTTC 719
Db 2772 TGTTTAAATTTATCGGTATATATATAAAGTTGAAATTTTCGTATAAAGTAGTTAT 2831
QY 720 AATACAAATTTTGGCTGATCTTGCACACTCTTCGATCATATCGCAATATACGTGAAC 779
Db 2832 TGTGAAATTTTAGTTGAATTCGAGATAGAAATTTTAAAGTTATTTTATTTT 2891
QY 780 ACTGGTGACTAAATTTGTTGTTAAATTTAGATTTCTATCTCCGGTTTAAAA 839
Db 2892 TTAGATTTTTTTTTTTTTTTTATTTATTAATTTGTTTAGTGAAATTTTAAATTTTA 2951
QY 840 GTGNAATTTATGATCATGCTTAAACATTTGTAAGTAAGATGATAATAAATGATAAAT 899
Db 2952 TAATATTAAAGTATAAGTAAATAAAGAAAAAGAGTTGGATTTTATTAATAAATTA 3011
QY 900 TAGTTG 905
Db 3012 AGTTG 3017

RESULT 14
US-08-894-731-1/c
; Sequence 1, Application US/08894731
; Patent No. 6084089
; GENERAL INFORMATION:
; APPLICANT: MINE, Toshiki
; APPLICANT: OHYAMA, Akio
; APPLICANT: HIYOSHI, Toru
; APPLICANT: KASAKA, Keisuke
; TITLE OF INVENTION: COLD-INDUCIBLE PROMOTER SEQUENCE
; FILE REFERENCE: 760-234P
; CURRENT APPLICATION NUMBER: US/08/894,731
; CURRENT FILING DATE: 1997-10-27
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3600
; TYPE: DNA
; ORGANISM: Solanum tuberosum
US-08-894-731-1

Query Match 3.8%; Score 43.6; DB 3; Length 3600;
Best Local Similarity 48.7%; Pred. No. 0.16;
Matches 148; Conservative 0; Mismatches 154; Indels 2; Gaps 1;

Search completed: October 9, 2004, 08:40:08
Job time : 148 secs

QY	658	CA	GTGTCACCTTGTTCTCTCGTGTCTTTTCTTTTAAATTTTTCATAAAGTCTCTTTTGTGTTTATCT	717
Db	2555	CA	TTTTCACAAATCGAATTCGACGTATTATATTTACTTATCGAATTCGAAATCTTTTTCAT	2496
QY	718	TC	CAATACAAATTTTTGGCGTGTATCTTGCAAACTCTTCGATCATATCGCCAAATATACG--T	775
Db	2495	TT	AAAAAAGTGAATGTGTGAATTTTTCATTTTCTTTTAAATGTGAATGTTCGAAG	2436
QY	776	GA	CACCTGGTGATCTAAATTTGTGTGTTAAATGTTTAAATTTAGATTTCTATTTCTCCGGTTT	835
Db	2435	TT	CCAAGAGAGTCTCAGAAAAATCTTTTCTTTTGTCTTTTATTTACATTTGCTAGGTG	2376
QY	836	AAA	AGTGAATTTATATGTAATCATGGTTAAACAATTGTAAGTAAGATGATAATAAAATGATA	895
Db	2375	AT	TAAAGAGATTTAAATTTTATTTATATTTATGAGTTTATTTGAAATTAATATATATTTTA	2316
QY	896	AA	TTTAGTTGATGATAACGCTGAAGCAAAAATCAGATAGATACATTTGATTTGTCGTA	955
Db	2315	AG	AGTAGTTTCAAAATTAACCTAAAGAGTAAAAATAAAAAATAATTTAAATTTATATCTA	2256
QY	956	TTTT	959	
Db	2255	ATTT	2252	
RESULT 15				
US-09-270-767-2481				
; Sequence 2481, Application US/09270767				
; Patent No. 6703491				
; GENERAL INFORMATION:				
; APPLICANT: Homburger et al.				
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster				
; FILE REFERENCE: File Reference: 7326-094				
; CURRENT APPLICATION NUMBER: US/09/270,767				
; CURRENT FILING DATE: 1999-03-17				
; NUMBER OF SEQ ID NOS: 62517				
; SOFTWARE: PatentIn Ver. 2.0				
; SEQ ID NO 2481				
; LENGTH: 1316				
; TYPE: DNA				
; ORGANISM: Drosophila melanogaster				
; FEATURE:				
; OTHER INFORMATION: n means any nucleotide				
US-09-270-767-2481				
Query Match 3.7%; Score 43.2; DB 4; Length 1316;				
Best Local Similarity 46.1%; Pred. No. 0.14;				
Matches 143; Conservative 0; Mismatches 163; Indels 4; Gaps 1;				
QY	672	TTT	CTCGTTTTTTTTTAAATTTTCTAAGTCTCTTTGTTTATCTTCAATACAAATTTT	731
Db	143	TTT	TNTNNNTATTTTNNANNTAAATAAATATTNGGGGGTTATAAAATATAAAAAATA	202
QY	732	TG	CGCTGATCTTGCAAACTCTTCGATCATATCGCCAATATACGTGAACACTGGTGATCTA	791
Db	203	TTT	TTTATTTTAAAAAATTTTTTATNTAATANNNAATAAAANTTAAANTTTTTTTTA	262
QY	792	ATT	TGTTGTTAAATG-----TTAAATTTAGATTTCTATTTCTCCGGTTTAAAAAGTGAATTA	847
Db	263	ATA	TATTTGGTTTATATGGAAATTAATANTATATTATTAATAATTAATANTTAATTA	322
QY	848	TAT	GATCATGCTTAAACAATTGTGAAGTAAGATGATAATAAAATGATAAAATTTAGTTGAT	907
Db	323	AA	TTTTTATTTGGTNAAAAATNTFAAATATATAAAATAAATAATAATAAATTAATTTAT	382
QY	908	GG	ATAACGTGAAGCAAAAAATCAGATAGATACATTTGATTTTGTGCGTATTTTGACATATG	967
Db	383	ATA	TGAGATAAATAAATTTATTTATGCGGATTTTNTGTAATGTTTTTTTTTTTTATTGAAGA	442
QY	968	CG	GAGAGTGA	977
Db	443	GT	GAGTGGGA	452

